Reduction of mathematical models of signal transduction networks: simulation-based approach applied to EGF receptor signalling

H. Conzelmann, J. Saez-Rodriguez, T. Sauter, E. Bullinger, F. Allgöwer and E.D. Gilles

Abstract: Biological systems and, in particular, cellular signal transduction pathways are characterised by their high complexity. Mathematical models describing these processes might be of great help to gain qualitative and, most importantly, quantitative knowledge about such complex systems. However, a detailed mathematical description of these systems leads to nearly unmanageably large models, especially when combining models of different signalling pathways to study cross-talk phenomena. Therefore, simplification of models becomes very important. Different methods are available for model reduction of biological models. Importantly, most of the common model reduction methods cannot be applied to cellular signal transduction pathways. Using as an example the epidermal growth factor (EGF) signalling pathway, we discuss how quantitative methods like system analysis and simulation studies can help to suitably reduce models and additionally give new insights into the signal transmission and processing of the cell.

1 Introduction

Cells consists of many chemical components interacting via biochemical reactions, allosteric effects, etc. For example, the 'simple' bacterium *Escherichia coli*, already possesses over 4400 genes, about 2500 different active proteins and enzymes, 50–70 sensors in the cell membrane, and a large number of pathways leading from substrates to intermediary products and cellular structures. This complexity hampers understanding of how cells function and what might be the reason behind their configuration. An important and complex part of a cell is its signal transduction system. As cells have to cope with very different environmental conditions, it is essential for them to continuously sense their environment and communicate with each other. For these purposes, cells employ a complex and robust signalling machinery, which is able to transmit and integrate a great number of different signals, controlling biological outcomes such as cell growth, differentiation, migration, and apoptosis.

The complexity of signalling phenomena is not intuitively understandable, and mathematical modelling is a promising tool for its analysis [1]. New experimental techniques, which provide an immense amount of data, have brought us into the position to tackle the challenge of modelling these signalling networks. However, the more our knowledge about these signalling systems is enhanced, the more the mathematical models increase in size and complexity. Hence, mathematical models, which should help us to understand the signalling processes, have become themselves too complex for intuitive understanding, making their analysis and interpretation a task almost as arduous as the original one, namely the analysis of the signalling processes.

Therefore, the simplification of the mathematical models is of great interest. Obviously, every model includes many implicit simplifications based on biological assumptions: neglect of parts of the cell which do not influence the phenomenon of interest, consideration of elements as constant pools, etc. However, a mathematical, systematic analysis of the models can provide additional simplifications, which might not be found in an intuitive manner. This theory-driven simplification of mathematical models is the objective of the field of model reduction.

The complexity of signalling phenomena arises in part from the amount of complexes that the molecules involved in signalling processes can build [2]. The receptors, and certain signalling molecules that act as scaffolds, usually offer a large number of different binding sites. As pointed out by Hlavacek *et al.* [2], the number of distinguishable species grows exponentially with the number of binding sites. For example, in the case of epidermal growth factor (EGF) receptor signalling there are about 1232 potential unique complexes of the EGF receptor species.

However, current models of signal transduction pathways only include a small part of all potentially existing species [2]. If we recalculate the number of species introducing the additional GAP binding (included for example in the model of Schoeberl *et al.* [3]), we find 4752 distinguishable chemical species (calculation in the Appendix Section 8). As the model of Schoeberl *et al.* also includes receptor internalisation the whole model would contain 9504

Systems Biology online no. 20045011

doi: 10.1049/sb:20045011

Paper first received 23rd March and in revised form 10th May 2004

H. Conzelmann and T. Sauter are with the Institute for System Dynamics and Control Engineering, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany, email: conzelmann@isr.uni-stuttgart.de

J. Saez-Rodriguez and E.D. Gilles are with the Max-Planck-Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany

E. Bullinger and F. Allgöwer are with the Institute for Systems Theory in Engineering, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany

E.D. Gilles is also at the Institute for System Dynamics and Control Engineering, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany

different species, neither including the unbound proteins like GAP, Grb2 and so on, nor the MAP kinase cascade. Remodelling the whole EGF signalling pathway introduced in [3] leads to a system consisting of 37 489 differential equations (calculation in the Appendix). It does not seem to be reasonable to create a model with nearly 40 000 ordinary differential equations (ODEs) to describe solely EGF signalling. Since several predictions of the model introduced by Schoeberl *et al.* could be validated, it does not seem necessary to consider all species. Therefore, the models like the one introduced by Schoeberl *et al.* are already intuitively simplified models.

However, intuitive model simplifications have a number of drawbacks. First, as already mentioned, it is not guaranteed that a minimal model, which may not be reduced further without losing essential information, will be found. Second, there is a lack of methodology which can be applied to other signal transduction systems. Therefore, it is desirable to perform model reduction in a systematic manner. We have developed a method to reduce models of signalling pathways that account for all chemical species, and it seems to be a promising method for model reduction in general (Conzelmann *et al.*, in preparation).

A promising concept in order to analyse biological systems and their models, potentially complementary to model reduction, is to dissect them into functional units [4, 5] and analyse the signalling networks in terms of the resulting modules [6]. However, even though biology is broadly considered modular, there is no generally accepted definition of the modules [7]. Different proposals, ranging from evolutionary conservation to graph-theory's clustering have been made [7]. An interesting criterion for the decomposition of biological networks into modules might be the absence of retroactivity [5]. Two modules A and B are connected without retroactivity if there is an influence from a submodule A1 from A to a submodule B1 from B, but the submodule B1 does not influence A1 directly. A feedback from B to A is possible, as long as it is not between the elements B1 and A1. Since the behaviour of retroactive-free modules only depends on inputs and initial conditions, their transfer performance can be studied uncoupled from the rest of the model [6]. Additionally, if model reduction is applied on modules free of retroactivity, the reduced model can be directly re-wired in the whole network.

In this contribution, first, the current methods available for the reduction of mathematical models of biological systems are briefly summarised. Second, a new approach, based on simulation studies and system analysis is proposed and applied to the model of Schoeberl *et al.* [3] for EGF receptor signalling.

2 Model reduction of signalling networks: state-of-the-art

Reducing the complexity of mathematical models is achievable by different model reduction methods. Classically, lumping is used for reducing biological models. Lumping refers to reducing the dimension by grouping states (metabolites) or flows. This approach is preferred by biologists as the states of the reduced system are either equal to the original ones or have a direct meaning, like sums of metabolites, for example. This is illustrated with the following example. In the following reaction scheme

$$\begin{aligned} S &\rightarrow I_1 \rightarrow P \\ S &\rightarrow I_2 \rightarrow P \end{aligned} \tag{1}$$

the species I_1 and I_2 can be lumped together to a virtual species $I = I_1 + I_2$, resulting in

$$S \rightarrow I \rightarrow P \tag{2}$$

The other two states, S, and P are present also in the reduced model. A second example are compartment models where, for example, all different proteins are regarded as a single protein compound. All different proteins are treated as one protein with an average composition.

Additionally, biological processes in general take place over a very broad timescale reaching from milliseconds (e.g. phosphorylation reactions) and seconds (e.g. metabolic conversions) to hours, days (e.g. gene expression) and years (e.g. differentiation, evolutionary processes). Model reduction based on these different timescales has been used, for example, to derive various kinetic rate expressions, the best known being the Michaelis Menten kinetic, where lumping is performed on a timescale-based model approach:

$$S + E \overset{v_1}{\rightleftharpoons} ES \overset{v_2}{\longrightarrow} E + P \tag{3}$$

Lumping the reactions v_1 and v_2 eliminates the intermediate ES and makes sense if the intermediate is in quasi steady state. The reduced model is then described by a single reaction whose rate depends on the enzyme and substrate concentration

$$S \overset{v(E,S)}{\longrightarrow} P \tag{4}$$

The Michaelis Menten equation, presented in 1925 by Briggs and Haldane, is of fundamental importance for enzyme kinetics. Although derived only for the simple case of an irreversible one-substrate reaction, it remains also valid, under certain conditions and modifications, for more complex mechanism [8, 9]. For a mathematical derivation of the Michealis Menten equation using invariant manifold methods, see [10].

The principle of timescale separation has been widely used for model reduction in biology. For examples, see [11–13]. Applying a more mathematical approach, invariant manifold methods were described for metabolic model reduction [10]. However, many processes in signal transduction pathways proceed in the same timescale (for example the EGF signalling discussed in this paper), and a model reduction based on the separation of different timescales becomes difficult and less fruitful.

In contrary to lumping and timescale-based reduction, model reduction algorithms from systems theory usually result in reduced models whose states are different from the original states. One method is to eliminate not well-observable states. This method requires the definition of a set of key variables (outputs of the signalling network), leading to a mathematical description of the form

$$\dot{\vec{x}} = \vec{f}(\vec{x}, \vec{u}, \vec{p}) \qquad \vec{x} \in \mathbf{R}^n, \vec{u} \in \mathbf{R}^m \tag{5}$$

$$\vec{y} = \vec{h}(\vec{x}) \qquad \vec{y} \in \mathbf{R}^r \tag{6}$$

where $\vec{x}$ are the system states, $\vec{u}$ are the inputs, $\vec{y}$ are the outputs and $\vec{p}$ are the parameters. The definition of outputs is essential, since observability is not a fixed property of a dynamical system, but dependent on the choice of the output variables $\vec{y}$.

To illustrate these ideas, consider the simple dynamical system

$$\dot{x}_1 = -x_1^2 \tag{7}$$

$$\dot{x}_2 = x_1(x_2 - 1) \tag{8}$$

Choosing the output $y_1 = x_1$, the dynamics of x_2 are not observable, since x_2 can not be calculated from x_1: changes

in x_2 would not affect x_1. However, the dynamics of both states can be observed if $y_2 = x_2$ is the output.

For a detailed description of the particular observability based methods 'balancing' and 'projection on the main subspace', two methods based on the concept of observability, see [14–16]. The main advantages of these methods are that they can be automatically applied to large systems and that the resulting approximation error can be quantified. One drawback is that the states of the reduced model are different from the original states and don't have a direct biological meaning. Another drawback is that some metabolite concentrations might become negative due to the approximation errors of the projections used in the algorithm. Therefore, the common model reduction methods which are based on observability cannot be applied successfully to cellular signal transduction pathways or to the modules of which they are defined.

This paper proposes a new model reduction approach by allowing lumping to be performed in a systematic way using a simulation based approach.

3 Methods

A new heuristic approach was developed combining several methods from systems engineering. It starts with the decomposition of the network into modules with low retroactivity. As mentioned earlier, a meaningful approach to analyse big non-linear mathematical models is to decompose them into smaller modules [4–5], as these modules often are more manageable and easier to analyse than the whole model. In order to reduce a model in size, the resulting non-linear modules shall be replaced by less complex non-linear models, showing approximately the same input/output (I/O) behaviour. This substitution requires a detailed knowledge of the model dynamics and its transfer behaviour, which not only helps to find suitable model reductions but also to get a qualitative and quantitative comprehension of the system. However, only few tools exist to analyse the dynamics of non-linear models. Additionally, most of the tools that can be found in the literature are restricted to the analysis of low-dimensional systems. One example is the determination of the steady states. Even for most low dimensional systems, this calculation can only be performed numerically. For high-dimensional problems the numerical calculations get time-consuming and are prone to errors.

Therefore, a detailed analysis of the system dynamics may be useful in order to find suitable model reduction methods that may allow systematic treatment of these systems. We will focus on the search for special patterns in the model structure and system behaviour (see also [17]), and investigate which dynamics are negligible in these patterns.

In our example, the steady states of each module can be calculated numerically, and the results show that each module has only one relevant equilibrium for the given parameter values $\vec{p}_i$ (data not shown). We additionally tested this result for different stationary input levels, and always found just one relevant steady state.

In contrast to most non-linear analysis tools, quantitative methods like simulation studies can be applied to the analysis of the I/O behaviour of essentially all systems. Considering the system's output responses to input signals like steps, oscillations or pulses may help to categorise the system's transfer performance, and to understand the signal processing of cells. Additionally, the system response to a pseudorandom binary sequence (PRBS) signal can tell us whether the system returns to its initial state as soon as the input signal vanishes. This is important to assure that the model is able to react to a restimulation.

A comparison of the responses of modules to different input signals with the responses of standard linear transfer elements can be performed. This analysis also reveals the extent of non-linearity of the modules [18]. If $\mathcal{U}$ is the set of considered input signals, and we find a linear system providing approximately the same I/O behaviour for all $\vec{u} \in \mathcal{U}$, the module can be approximated by a linear model for these inputs. This method allows us to test whether the system may be regarded as approximately linear for a certain class of inputs, or what kind of non-linear behaviour the module shows for other inputs.

If we are interested in the I/O behaviour at different operating points, a state-space linearisation of the module i

$$\dot{\vec{x}}_i = A_i\vec{x}_i + B_i\vec{u}_i \tag{9}$$

$$\vec{y}_i = C_i\vec{x}_i \tag{10}$$

at these operating points provides additional information, and may also help to understand non-linear phenomena. The state-space linearisations have been derived in an analytical manner using Mathematica.

Besides the state-space description in the time domain, linear models can also be transformed into the frequency domain using the Laplace transformation. The resulting transfer functions $G_i(s)$ (which are defined as $\vec{y}_i = G_i(s)\vec{u}_i$) can be calculated by

$$G_i(s) = C_i(sI - A_i)^{-1}B_i \tag{11}$$

using the system matrices from 9 and 10. These transfer functions are packed descriptions of the I/O behaviour of a linear system. In linear control theory, transfer functions are widely used, since their simple structure allows an easy analysis of the system's transfer performance. For this reason, we will also use this description to analyse the transfer behaviour of linearised modules and to compare them with linear transfer elements. Such a comparison with linear transfer elements helps us to understand the purpose of a certain system structure in the analysed signal transduction network.

The standard transfer elements we consider here are temporally lagged proportional elements, derivative and integral elements, and combinations of this elements. A temporally lagged element (e.g. a PT_1 element) is given by

$$T\dot{y} + y = Ku \qquad G(s) = \frac{K}{Ts+1} \tag{12}$$

The systems output response y to an input signal u can be derived by solving the differential equation on the left side, or by evaluating $y = G(s)u$. The parameter T is a time constant determining how quickly the system respond to an input u, and K is called the amplification magnitude, determining whether the input signal is amplified or attenuated. Derivative elements (e.g. a DT_2 element) are given by

$$T_1T_2\ddot{y} + (T_1+T_2)\dot{y} + y = K\dot{u} \qquad G(s) = \frac{Ks}{(T_1s+1)(T_2s+1)} \tag{13}$$

and integral elements (e.g. a IT_1 element) are given by

$$T\dot{y} + y = K\int_0^t u(\tau)\mathrm{d}\tau \qquad G(s) = \frac{K}{s(Ts+1)} \tag{14}$$

Additionally, we also consider combinations of these standard elements, like e.g. a PDT_2 element

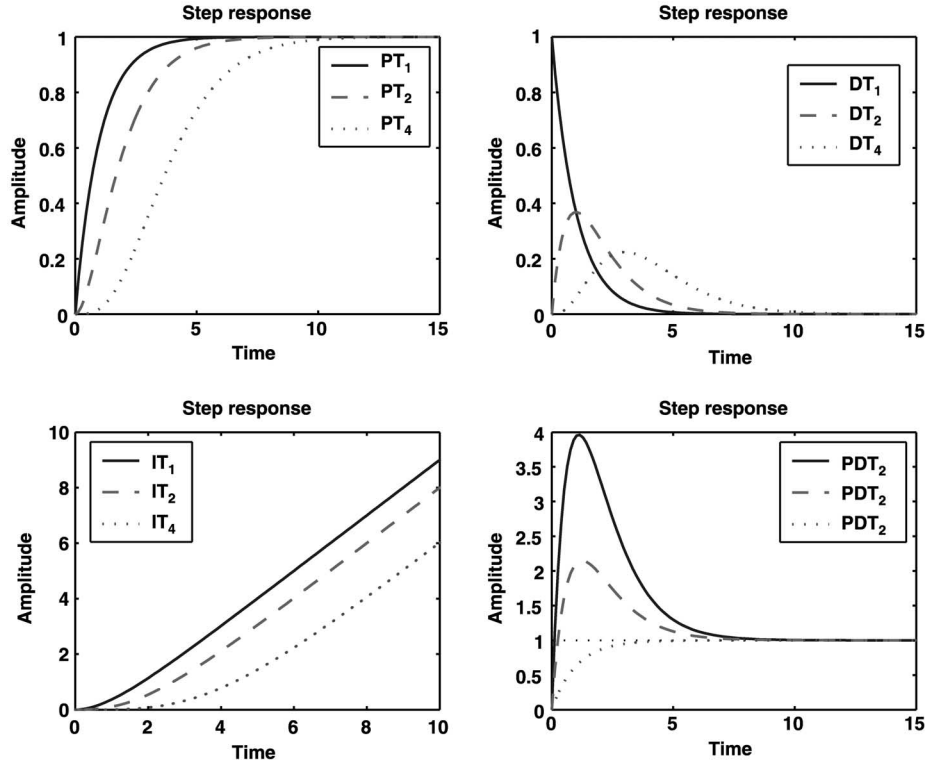

Fig. 1 *Step response curves of temporally delayed proportional elements, derivative elements and integral elements with different temporal delays. The step responses of the PDT_2 elements are characterised by differently strong derivative parts*

$$T_1 T_2 \ddot{y} + (T_1 + T_2)\dot{y} + y = K(\dot{u} + u)$$
$$G(s) = \frac{K(s+1)}{(T_1 s + 1)(T_2 s + 1)} \qquad (15)$$

Typical step responses of these linear systems are shown in Fig. 1.

In summary, the new heuristic approach combines several known tools. It starts with the decomposition of the network into modules with low retroactivity. Simulation studies and structural comparison of the modules reveal common patterns. These patterns are formulated with focus on observable groups of states. In contrast to other observability-based model reduction methods from systems theory, the biological meaning of the states of the reduced model is mostly preserved in our approach.

4 Case study: EGF receptor signalling

As a case study, we have investigated the mathematical model of the EGFR induced MAPK cascade of Schoeberl *et al.* [3]. The EGF signalling network is probably the best known signalling network in mammalian cells. Importantly, modelling has helped to improve our knowledge of EGF-related signalling phenomena [19]. The EGF receptor can bind several ligands like EGF and TGFα [20]. The binding of ligand facilitates the dimerisation of the EGF receptors. The dimerised EGFRs phosphorylate reciprocally, creating docking sites for signalling molecules, and phosphorylate signalling molecules [20], activating many signalling pathways. One of these pathways is the Raf/MEK/ERK MAP Kinase cascade, which is activated through two pathways, one Shc-dependent and one Shc-independent.

The model [3] includes all the steps which link the binding of ligand to receptor to activation of ERK: (i) ligand binding and EGF receptor activation, (ii) formation of signalling complexes, (iii) activation of Ras and the consequent activation of the Raf/MEK/ERK MAP Kinase cascade. The model also includes the internalisation of the receptors.

We have previously decomposed the model according to the criterion of absence of retroactivity, obtaining modules which are connected in most of the cases with a weak retroactivity [5] (see Fig. 2). The modules correspond roughly to the biological steps described above: EGF reception (1), complex formation (2 and 3), the activation of Raf and Ras (which are strongly coupled and therefore considered as a single unit)(4), the activation of MEK (5) and the activation of ERK (6). The complex formation module can be decomposed in two submodules, which correspond to the pathways independent (2) and dependent (3) of Shc. These two submodules are strongly coupled [5].

Having decomposed the EGF signalling network, simulation studies for each module can be performed. Simulations were performed using Matlab. Figure 3 shows characteristic step response curves of each module, which are discussed next.

4.1 MAP kinase cascade

First, we shall discuss briefly the MAP kinase cascade, which has already been extensively investigated in the past from a system-theory point of view (e.g. [21, 22]). Figure 3*e* shows the step response of the MAP kinase kinase module (MEK), and Fig. 3*f* the step response of the MAP kinase module (ERK). Both modules are equally structured but differ in their parameter values. Simulation studies indicate that both modules qualitatively show the same I/O behaviour. Figure 3*e* implies that this module approximately behaves like a temporally delayed integral element until the saturation point is reached. This can be shown by rewriting the model equations using Michaelis Menten kinetics. Considering the resulting differential equation for [MEK-PP]

$$\frac{d[\text{MEK-PP}]}{dt} = \frac{k_1[\text{MEK-P}]}{K_{m,1} + [\text{MEK-P}]}[\text{Raf}^*] - \frac{k_2[\text{MEK-PP}]}{K_{m,2} + [\text{MEK-PP}]}, \qquad (16)$$

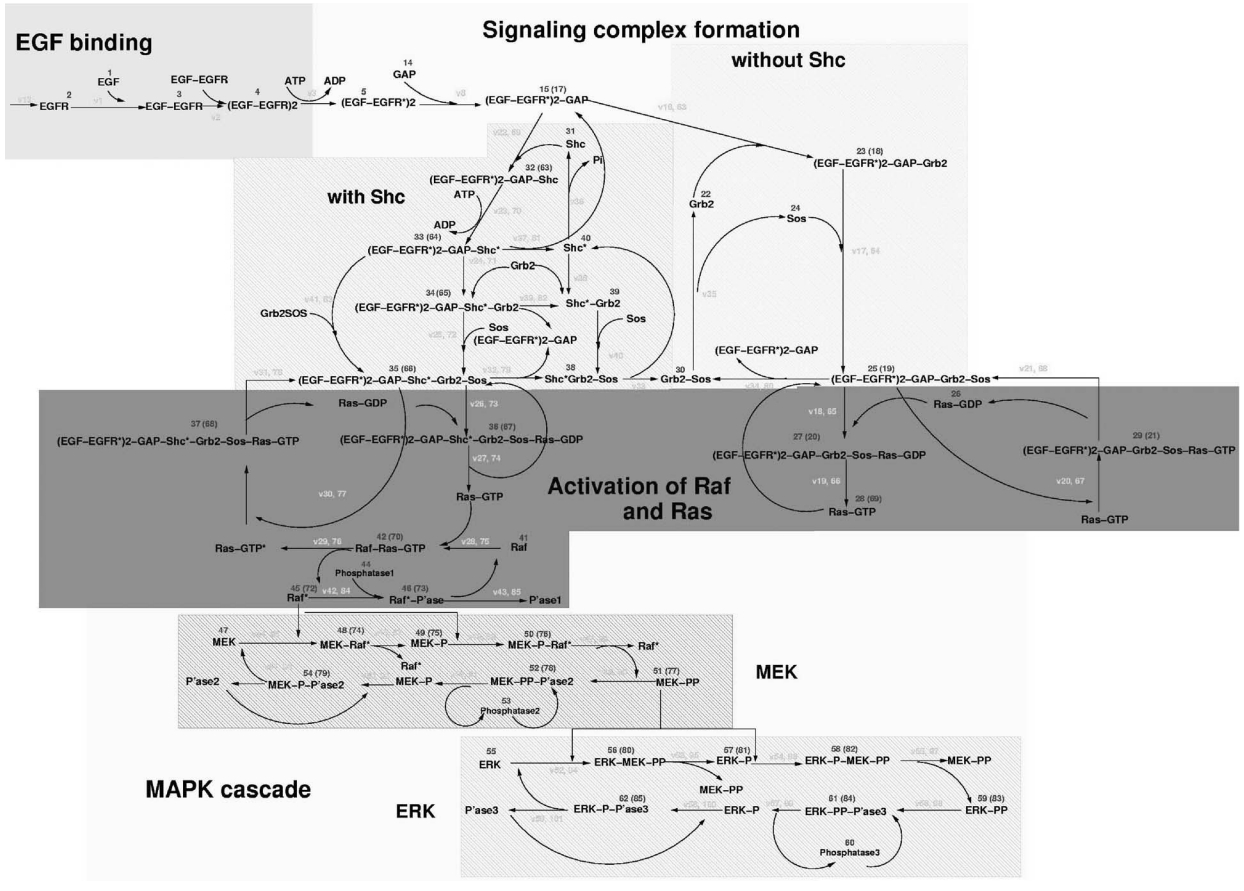

Fig. 2 *Modular decomposition of the reaction schema of the model for the EGF induced MAPK Cascade of Schoeberl et al. [20] (adapted from Schoeberl et al. [20] with the permission of the author)*

the transfer performance of this module can be analysed easily. Assuming $[\text{MEK-P}] \gg K_{m,1}$ and $[\text{MEK-PP}] \gg K_{m,2}$ (which is given for high input concentrations), the two fractions in (16) can be approximated by k_1 and k_2 and the differential equation

$$\frac{d[\text{MEK-PP}]}{dt} = k_1[\text{Raf}^*] - k_2 \tag{17}$$

is very similar to a linear integral element. Assuming $[\text{MEK-P}] \ll K_{m,1}$ and $[\text{MEK-PP}] \ll K_{m,2}$ (which is given for low input concentrations) and that $k_1[\text{MEK-P}] = k(t)$, we approximately get a differential equation realising a temporally delayed proportional element with a time variant attenuation $k(t)$ [6].

4.2 Analysis of the EGF receptor binding

The output curve of the EGF reception module (the concentration of $(\text{EGF-EGFR})_2$ see Fig. 3*a*) is similar to the step response of a DT_2 element with a weak proportional part. However, this analogy does not hold for vanishing input signals, since the response curve of a linear DT_2 element to such an input would run into the negative half-plane (equivalent to negative concentrations). This problem is discussed considering a linearisation of the module. To enhance the readability we assume that all reactions in the EGF receptor binding module are irreversible, which we have found to be a reasonable approximation (data not shown). We linearise the receptor activation module, including the reactions

$$\text{EGF} + \text{EGFR} \xrightarrow{k_1} \text{EGF-EGFR} \tag{18}$$

$$\text{EGF-EGFR} + \text{EGF-EGFR} \xrightarrow{k_2} (\text{EGF-EGFR})_2 \tag{19}$$

and

$$(\text{EGF-EGFR})_2 \xrightarrow{k_3} (\text{EGF-EGFR}^*)_2 \tag{20}$$

at the equilibrium point $\vec{x}_s = ([\text{EGFR}]_s, [\text{EGF-EGFR}]_s, [(\text{EGF-EGFR})_2]_s)$. The concentration of $(\text{EGF-EGFR}^*)_2$ is chosen as output y_1 of this module. The resulting system matrices A_1, B_1 and C_1 are transformed into the transfer function

$$G(s) = \frac{2[\text{EGFR}]_s[\text{EGF-EGFR}]_s k_1 k_2 s}{(s + k_1[\text{EGF}])(s + 2k_2[\text{EGF-EGFR}]_s)(s + k_3)} \tag{21}$$

using (11). Comparing this transfer function with (13), it can be found that the system behaves like a temporally delayed derivative element (DT_3) at operating points with $[\text{EGF}]_s > 0$. However, when the EGF input signal vanishes completely, the term $(s + k_1[\text{EGF}]) = s$ can be canceled with the s in the numerator. The remaining transfer function is a typical temporally delayed proportional element (PT_2). If the EGF concentration is not zero but the concentration of [EGFR] or [EGF-EGFR] is approximately zero, the term in the numerator and therefore the output signal also is approximately zero.

The analysis shows that the module works like a linear DT_3 element, whose parameters depend on the operating point. The most crucial non-linear effect is the limitation of the output signal. Hence, the structural pattern of this module (see Fig. 4*a*, further on called pattern A) defines a system designed for detecting changes in the input signal. This means that the cell can react quickly to changes in the input signal, which is important to guarantee a fast response to a stimulus.

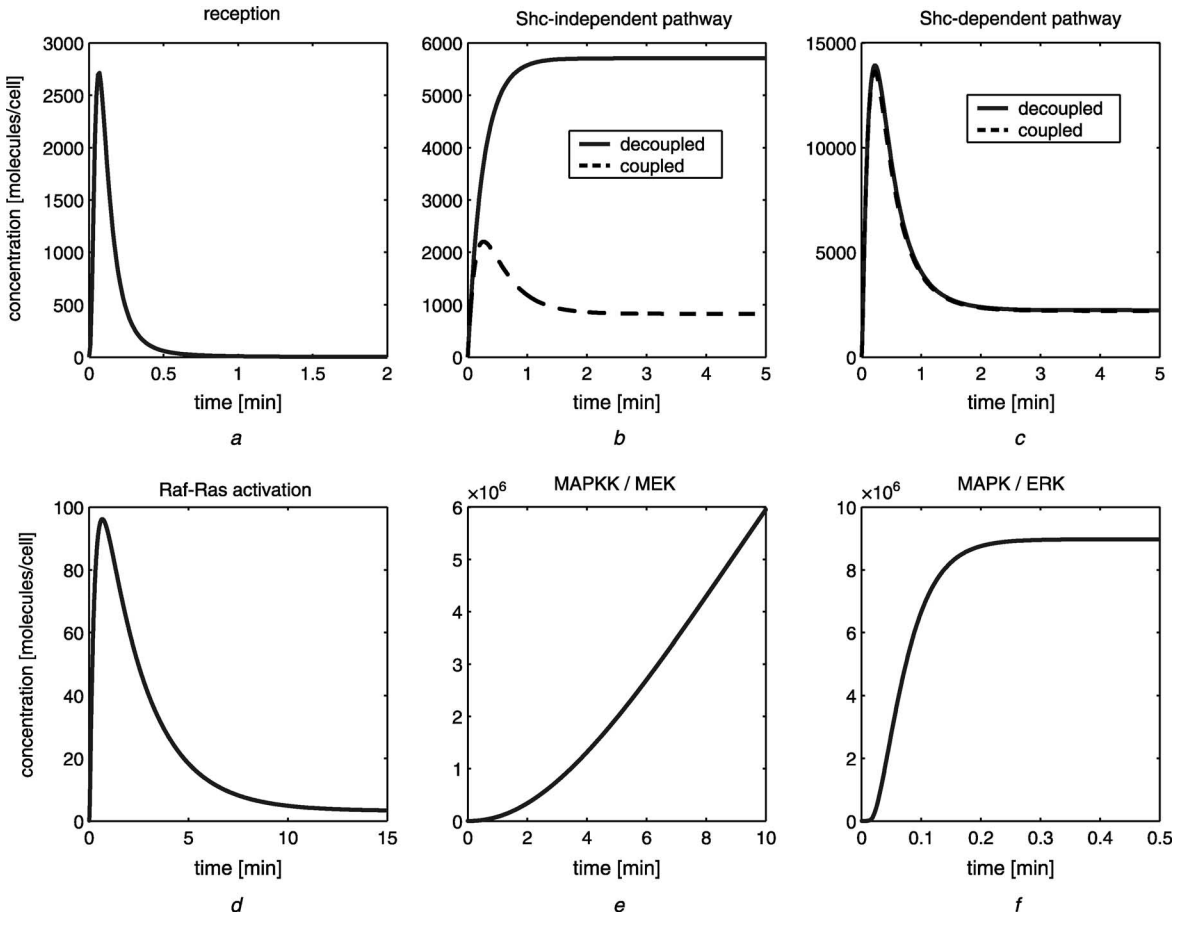

Fig. 3 *Step response curves of modules of EGF receptor signalling. The time windows are adjusted to the time constants of the modules*

a Input signal u_1 of the reception module is a step in the EGF concentration (from 0 to 8 nM). Output signal y_1 is the concentration of $(\text{EGF-EGFR})_2$

b Input signal u_2 to the Shc-independent pathway is a step in the $(\text{EGF-EGFR}^*)_2$-GAP concentration (from 0 to 1000 molecules/cell). Output signal is the concentration of $(\text{EGF-EGFR}^*)_2$-GAP-Grb2-Sos. The solid curve is the step response of this pathway, if the Shc-dependent pathway is disabled, while the dashed one is the step response if both pathways are working simultaneously

c Same boundary conditions like plot *b*. Output signal is the concentration of $(\text{EGF-EGFR}^*)_2$-GAP-Shc*-Grb2-Sos. The solid curve represents the response of the Shc-dependent pathway if the Shc-independent pathway is disabled, and the dashed one if both the Shc-dependent and Shc-independent pathways are active

d Input signal u_3 of the Raf-Ras module is a step in the output concentration of the Shc-dependent or Shc-independent pathway (from 0 to 1000 molecules/cell). Activated Raf* is the output signal of this module

e The input signal of the MAPKK module is a step in the Raf* concentration (from 0 to 100 molecules/cell), while the output signal is the concentration of MEK-PP

f The input signal of the MAPK module is a step in the MEK-PP concentration (from 0 to 10000 molecules/cell). The output signal is the concentration of ERK-PP. Note that plot e has the same form as the first part of the plot f, since both have a similar behaviour but a different timescale (see text)

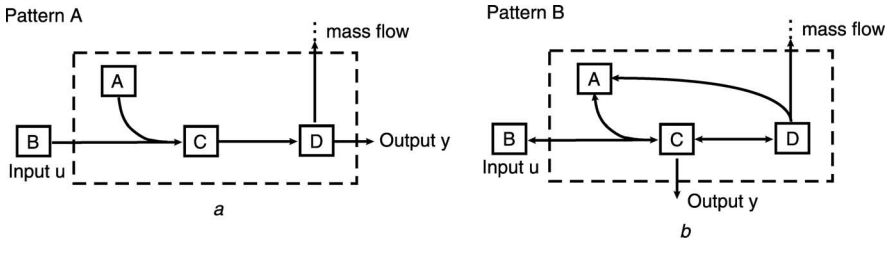

Fig. 4 *Pattern found in the EGF signalling cascade. Pattern A: A schematic representation of the reception module. The input B corresponds to the concentration of EGF, the pool of EGF receptors EGFR is represented by A, the resulting complex EGF-EGFR corresponds to C and the dimeric receptor complex* $(\text{EGF-EGFR})_2$ *is the output concentration D. Pattern B: A schematic representation of the Shc-dependent and the Shc-independent pathway. Input signal (B) is the concentration of* $(\text{EGF-EGFR}^*)_2$*-GAP. A represents the combined pools of Grb2, Sos and Shc. C corresponds to the complex* $(\text{EGF-EGFR}^*)_2$*-GAP-Shc***-Grb2-Sos or* $(\text{EGF-EGFR}^*)_2$*-GAP-Grb2-Sos, which can dissociate again (the effectors (A) are recycled)*

4.3 Complex formation (Shc-dependent and Shc-independent pathway)

The module describing the complex formation is the most complicated one. It includes two parallel signal transduction pathways (the Shc-dependent and the Shc-independent), which are highly coupled by the effector molecules Grb2 and Sos. Additionally, both pathways are coupled with the binding of GAP, and there is no possibility to separate them with respect to the criterion of absence of retroactivity. Nevertheless, a separate analysis is still reasonable, as we are interested in the similarities and differences of these pathways. First, we start with typical step responses of the two pathways (see Figs. 3*b* and 3*c*). The Shc-dependent pathway shows a strong peak overshoot being evocative of

a *PDT* element. Interestingly, the I/O behaviour of this module scarcely changes if it is coupled with the Shc-independent pathway. The step response of the Shc-independent pathway is evocative of a *PT* element, and is strongly influenced by the coupling to the Shc-dependent pathway.

Although the system structure of both pathways is quite similar, they show a very different I/O behaviour. Interestingly, the dynamics of both modules can be reproduced by a simple system which we shall call pattern B (see Fig. 4*b*). The only difference between pattern A and B is that the molecule A is recycled. Comparing pattern B to the complex formation module, the effectors Shc, Grb2 and Sos would be lumped in A. Assuming that the pooling of the effector molecules does not affect the principal I/O behaviour, it must be possible to ascribe these differences to the model parameters. The core structure of these modules is hence pattern B, which is defined by the following set of reactions

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \tag{22}$$

$$C \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} D \tag{23}$$

$$D \xrightarrow{k_3} A \tag{24}$$

In the following we will show that the differences in the I/O behaviour can be explained using this simplified model. The only non-linearity in this system is the dependence of the forward reaction 22 on the product $c_A c_B$. As this is the product of a state with the input, we can interpret this model as a linear time-variant one. If the system is stimulated by an initial step in the concentration c_B at $t = 0$, we only have to deal with a linear time-invariant system. As a step in the concentration is not physiological, we also considered signals with a slower rise in the input concentration (see Fig. 5). However, one always has to keep in mind that there may also be input signals for which the non-linearity will influence the I/O behaviour significantly. The transfer function of the linearised model is a PDT_2 element. Now we want to show how the step response of a PDT_2 element

$$G(s) = \frac{k(s + a_0)}{(s + a_1)(s + a_2)} \tag{25}$$

depends on the parameters. In the reactions scheme (22) to (24), the parameters k, a_0, a_1 and a_2 can be identified very easily by linearising the model and transforming the state space equations to a transfer function (like shown above). In the following, especially the parameters $k = k_1 c_{A,s}$ and $a_0 = (k_{-2} + k_3)$ are important (the equivalences of parameters a_1 and a_2 are included in the Appendix). The choice of the parameter a_0 defines, for example, whether the proportional or the derivative element dominates the system performance ($a_0 = 0$ equals a DT_2 system, $a_0 \gg 1$ approximates a PT_2). It is useful to know whether a system has a peak overshoot (i.e. a maximum) or not, because if the step response of a linear system has a peak overshoot and does not oscillate (comparable to the step response of the Shc-dependent pathway), the system has a derivative component. Importantly, the lack of a peak overshoot does not necessarily imply the lack of a derivative element. The analytical time function of the step response is given by the equation

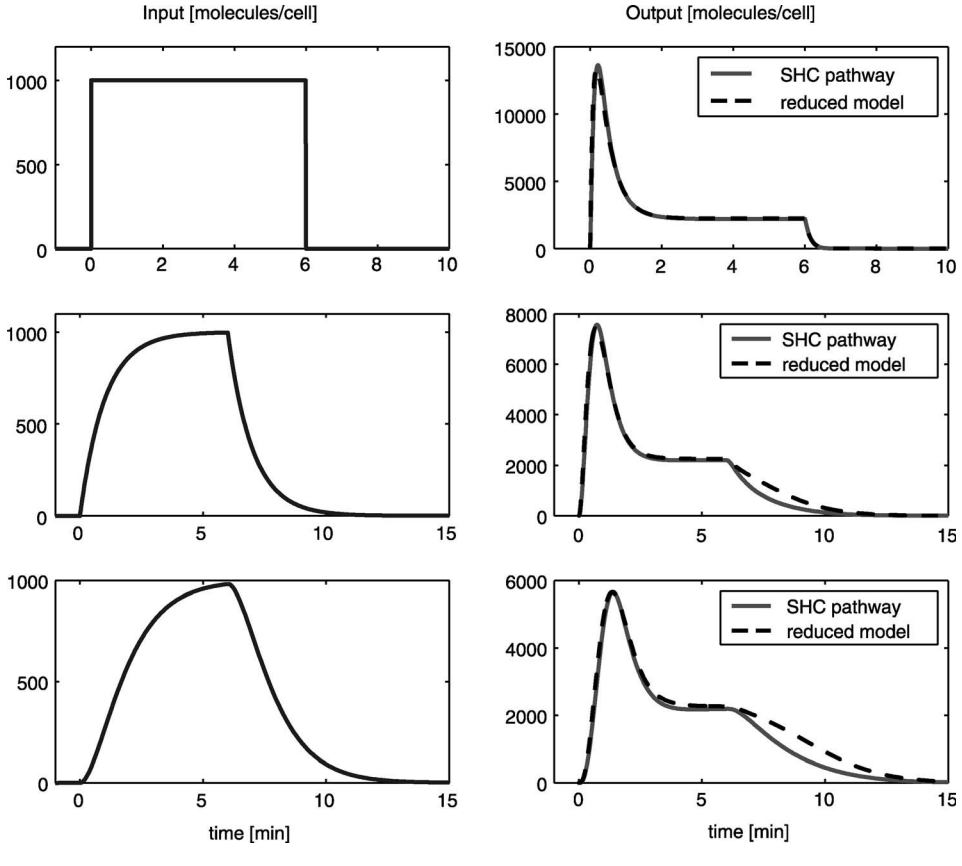

Fig. 5 *The reduced model approximates the I/O behaviour of the Shc-dependent pathway very well, even for different input signals. However, the match of the stationary final value gets worse if the amplitude of the input signal varies highly from the one chosen here (data not shown). Since the module is supposed to have a derivative element, its response to input signals with different initial gradients is of special interest*

$$y(t) = k\left(\frac{a_0}{a_1 a_2} + \frac{a_0 - a_1}{a_1(a_1 - a_2)} e^{-a_1 t} + \frac{a_2 - a_0}{a_2(a_1 - a_2)} e^{-a_2 t}\right) \tag{26}$$

Assuming that a_1 and a_2 are real numbers, $y(t)$ has zero or one peak overshoot for all $t \in \mathbf{R}_0^+$. Since the maximum is reached at

$$t_0 = \frac{ln\left(\frac{a_0 - a_1}{a_0 - a_2}\right)}{a_1 - a_2}, \tag{27}$$

the step response has a peak overshoot if $a_0 < a_1$ and $a_0 < a_2$. If these two conditions are not fulfilled (27) does not have positive real solutions. Analysing the parameters a_0, a_1 and a_2 (see Appendix), we find that for high recycling rates of the compound A (high value for k_3), the system behaves more like a proportional element, while its derivative part is stronger for lower recycling rates. The initial gradient of the curve ($\dot{y}(t = 0) = k$) can give us information about how quickly the output signal is rising. Thus, the concentration c_A determines how fast the system responds to an input signal (since $k = k_1 c_A$) and also how high this response is (since $y(t)$ is proportional to k, see (26)).

Additional information can be deduced from the stationary final value of the output concentration

$$c_{c,s} = \frac{k_1\, c_{A,s} c_{B,s}(k_3 + k_{-2})}{k_2\, k_3 + k_{-1}(k_3 + k_{-2}) + c_{B,s} k_1 (k_2 + k_3 + k_{-2})} \tag{28}$$

(28) indicates that for low input signals the stationary output concentration nearly is proportional to the input concentration c_B, but is saturated for high input signals.

The analysis of the relatively simple pattern B facilitates the interpretation of the results from the simulation studies of more complex modules, such as the Shc-dependent pathway. The amplification factor as well as the initial gradient of the systems step responses are influenced by the concentration of the effectors Shc, Grb2 and Sos (data not shown). Furthermore, the characteristics of the stationary I/O behaviour of both systems also correlate with the performance of the discussed pattern B. For low input signals the stationary output concentration is proportional to the input, but saturates for high input signals.

4.4 Raf-Ras activation

Interestingly, the structure of the Raf-Ras module (involving the activation of Ras-GDP and Raf) can also be compared with pattern B. Here the input signal (the concentration of the receptor complexes bound to Sos) interacts with Ras-GDP and thereafter with Raf to produce the output signal (activated Raf^*). Ras-GDP as well as Raf is recycled in this module. Figure 3*d* shows the step response of the Raf-Ras module, which also fits into pattern B and seems to be the realisation of a temporally delayed derivative element. The high number of modules in the EGF signalling pathway that resemble pattern B indicates the importance of this pattern. However, the transfer performance of the Raf-Ras module and pattern B differ considering a restimulation after the first stimulus vanished. While pattern B returns to its initial state, as soon as the input signal vanishes, the Raf-Ras module is not capable of returning to its initial state. For this reason, this module can only be replaced by pattern B for non-recurring input stimuli.

4.5 Model reduction

The analysis of the EGF pathway in the previous subsections shows that several modules are similarly structured and their structure as well as their qualitative

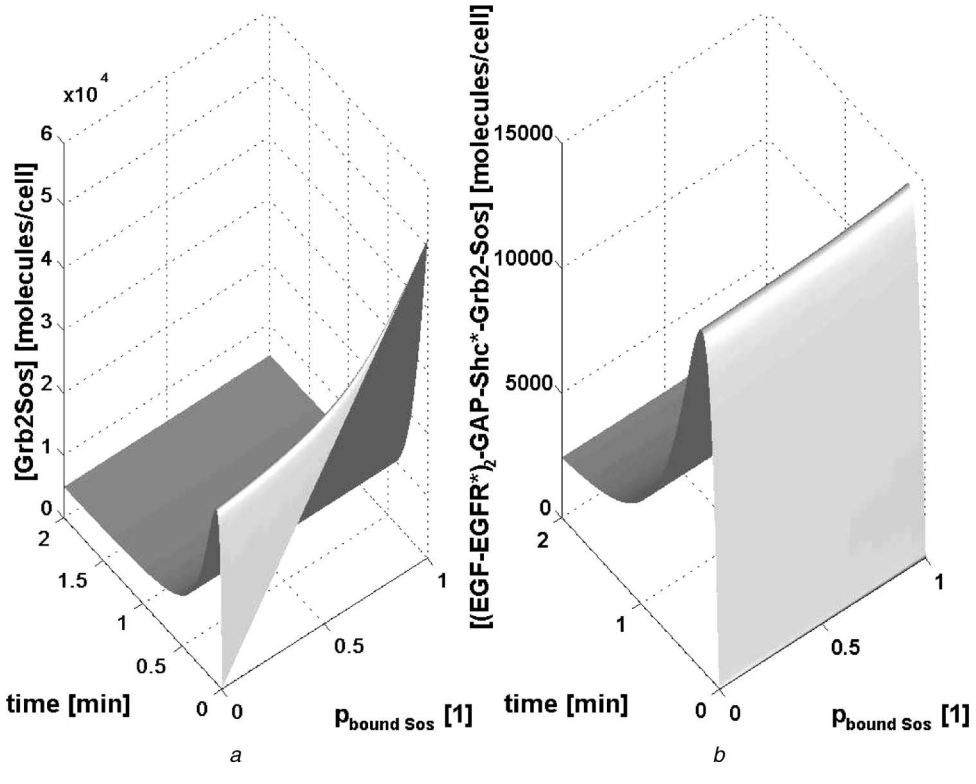

Fig. 6 *Time response of the Shc-dependent pathway for different initial conditions. a shows the concentration of the complex Grb2Sos. For all simulations the total concentration of both Grb2 and Sos is constant. The percentage p of Sos bound to Grb2 is changed from zero to one. Plot a shows that the differences in the starting point subside very quickly, meaning that all trajectories converge to a manifold, or low dimensional subspace. Additionally, plot b shows that even the differences in the very early times can hardly be observed at the output concentration* $(EGF\text{-}EGFR^*)_2$*-GAP-Shc*-Grb2-Sos*

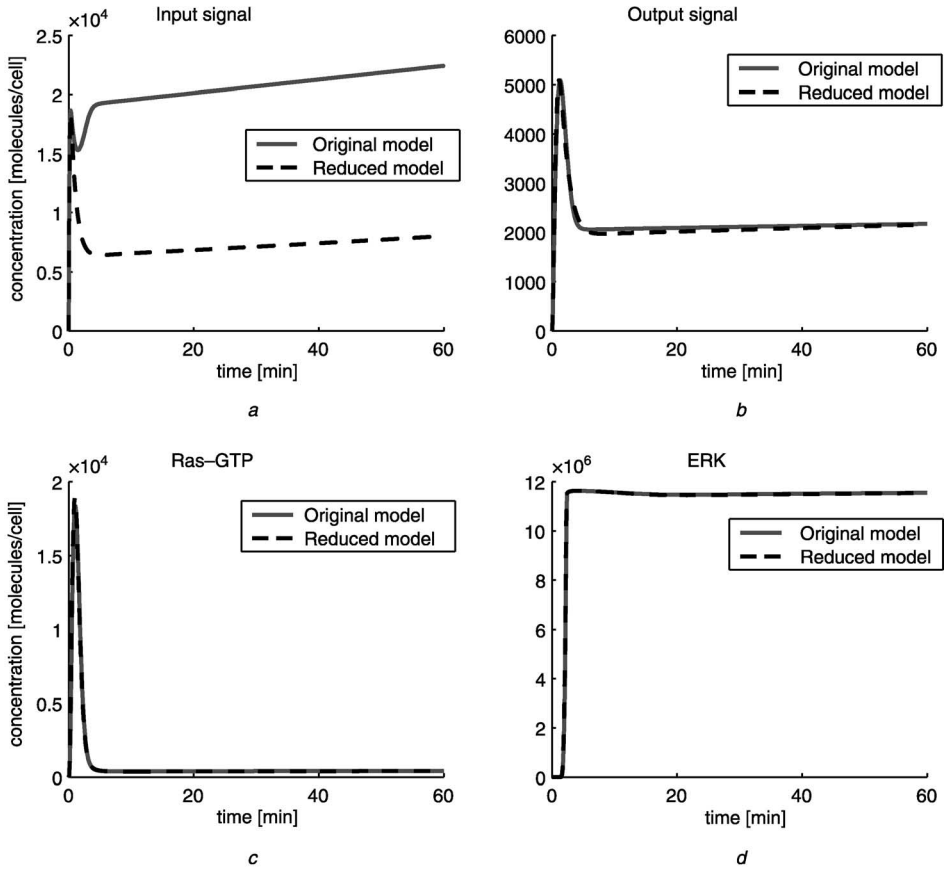

Fig. 7 *Comparison of the original EGF model by Schoeberl et al. (without receptor internalisation) and a model where the module for complex formation has been replaced by a model of pattern B. In both cases the system is stimulated by a step in the EGF concentration*

a shows the concentration of $(\text{EGF-EGFR}^*)_2$ which is the input signal of the module for complex formation and the output signal of the module for receptor binding. The run of this concentration is very different in the original and the reduced model. This shows that the weak retroactive effects between the module describing the receptor binding and the one describing the complex formation may already be problematic if a module in the system is replaced by a reduced version

b Output signal of the complex formation module ($(\text{EGF-EGFR}^*)_2$-GAP-Shc*-Grb2-Sos + $(\text{EGF-EGFR}^*)_2$-GAP-Grb2-Sos) and c_c of the original and reduced model, respectively

c Ras-GTP concentration in both models

d ERK-PP concentration in both models

I/O behaviour can be approximated by pattern B. Now, we want to discuss whether pattern B is also capable of approximating the quantitative system behaviour of these modules, and therefore can be used for model reduction. The Shc-dependent pathway was chosen as an example.

In pattern B the effectors Shc, Grb2 and Sos are lumped together in A. If this pattern is to be able to approximate the real I/O behaviour of the module, it must not be affected by the different dynamics of the effectors. This means that the dynamics of these states have to be either very fast or non-observable (see Section 2). Figure 6a implies that there exists a low dimensional subspace (or manifold) to which all curves converge quickly. The different runs of the curves at early times are nearly indistinguishable in the output concentration (Fig. 6*b*). Figure 5 shows that the I/O behaviour of the reduced system approximates the original system behaviour excellently. Therefore, the suggested model reduction seems to be appropriate.

However, this model reduction is only useful if the reduced model can be reintegrated into the whole EGF signalling model. When the EGF signalling pathway was decomposed, modules characterised by the absence of retroactivity were chosen. However, none of the modules is completely free of retroactive effects, what is notably true for the Shc-dependent pathway, which is highly coupled with the Shc-independent pathway. Additionally, there are also bidirectional couplings to the EGF receptor binding and the Raf-Ras module. As shown by Gong and Zhao [23] the Shc-dependent pathway is the more dominant one, what can be substantiated by sensitivity analysis. For this reason, we assume that our reduced model (model describing pattern B) will be able to approximate the effects of both pathways at once. To minimise the problems caused by bidirectional couplings the whole second module (GAP-binding, Shc-dependent and Shc-independent pathway, see Fig. 2) was replaced by the reduced model. Simulation studies imply that this larger module still can be approximated by the reduced model.

The model parameters have to be readjusted, as pattern B replaces the whole second module and not only the Shc-dependent pathway. However, it is possible to show that the reduced model represents a useful model reduction. In Fig. 7 different concentrations of the original model and the reduced model are compared. The runs of the concentration curves of ERK-PP at the end of the signalling cascade as well as the curves of the Ras-GTP concentration are nearly indistinguishable. Although the reduced module is coupled to the Raf-Ras module, the output signal of the reduced module and the original one differ only slightly. In contrast to these small differences, the input signals (the concentration of the dimerised and unphosphorylated receptor $(\text{EGF-EGFR})_2$) diverge in the two models. This discrepancy

is due to the retroactive effects which were neglected when the model was decomposed. Although these retroactive effects are very small compared to other couplings in the system, they lead to quite large differences in the $(EGF\text{-}EGFR)_2$ concentration. However, the output of the replaced module and, most importantly, the output of the whole system (concentration of ERK-PP) are remarkably similar for the reduced and the original model. Thus, pattern B can be used as a model reduction, which excellently approximates the I/O behaviour of the EGF signalling pathway.

5 Conclusions

As the immense complexity of biological systems is not intuitively understandable, mathematical modelling is a helpful tool in order to analyse the qualitative and quantitative properties of these systems. Since the size of the models is increasing, model reduction becomes more and more important to keep the models manageable. We have presented a brief survey of different model reduction methods with emphasis on their applicability to models of biological systems, especially to signal transduction. The modelling of signal transduction pathways is quite challenging, since, among other reasons, the number of distinguishable species is immense [2], and additionally most of the common model reduction methods are not straightforwardly applicable to these systems.

Usually this complexity is partially avoided in the modelling process by making explicit or implicit assumptions simplifying the model and hence reducing its size as, for example, in the EGF model by Schoeberl *et al.* [3]. We have taken this model as a case study and shown how quantitative methods like simulation studies may help to provide greater insight into the cell signal processing mechanism and to further reduce the model. For these purposes, the model was decomposed into smaller modules, that can be analysed more easily. The modules were chosen with weak retroactivity, according to the criteria introduced elsewhere [5], since this implies that the analysis and model reduction performed on the isolated modules can be applied to the whole network.

Interestingly, we have found a recurring pattern (pattern B shown in Fig. 4*b*), which is able to reproduce the dynamics of several subsystems. A deeper analysis of this pattern shows that for various input signals it is comparable to a linear *PDT* element (data not shown), enabling the cell to deduce information about the amount of EGF in the environment and additionally about the rate of change of this concentration. The parameters determine whether the proportional or the derivative element prevails. The information deduced from this analysis additionally helps to reduce the whole EGF model. For example, the module describing the complex formation can be replaced by pattern B. Figure 7 shows that the reduced model excellently approximates the I/O behaviour of the whole pathway. The modules which can be approximated by pattern B involve different biochemical processes. For example, the Raf-Ras module is mainly a transfer of phosphate between two proteins, while the complex formation modules involve mainly physical interactions among proteins. However, they can be both reproduced by pattern B, meaning that the dynamics, at least as far as we have analysed them, are very similar. It is tempting to speculate whether pattern B might also be applied to other signal transduction systems; it might be that, due to its characteristics for the transfer of information, it is a recurring theme in signal transduction.

In summary, we have shown that quantitative methods like simulation studies might be very helpful to analyse the I/O behaviour of a system, and may also be helpful in finding suitable model reductions. However, these methods might not be applicable to all models, as there may exist models that cannot be decomposed into retroactive-free modules, and, as shown by Fig. 7*a* the reduced modules cannot substitute perfectly the original one if the analysed units are not completely free of retroactive effects. Furthermore, this approach is rather heuristic, since it is case-dependent and requires a certain expertise from the analyser (e.g. for finding out a suitable reduced structure). The general applicability of this modus operandi is an open question. This method should be applied to other systems, in order to determine the range of applicability of this simulation-based approach.

We also expect the development of further methods for model reduction which, as discussed above, is a complicated task due to the idiosyncrasy of signal transduction systems. As usually happens in systems biology, biology challenges the abilities of systems theory.

6 Acknowledgments

We thank D. Zak for carefully reading the manuscript.

7 References

1 Asthagiri, A.R., and Lauffenburger, D.A.: 'Bioengineering models of cell signaling', *Annu. Rev. Biomed. Eng.*, 2000, **2**, pp. 31–53
2 Hlavacek, W.S., Faeder, J.R., Blinov, M.L., Perelson, A.S., and Goldstein, B.: 'The complexity of complexes in signal transduction', *Biotechnol. Bioeng.*, 2004, **84**, (7), pp. 783–794
3 Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., and Müller, G.: 'Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors', *Nat. Biotechnol.*, 2002, **20**, (4), pp. 370–375
4 Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W.: 'From molecular to modular cell biology', *Nature*, 1999, **402**, (6761-supp), pp. C47–C52
5 Saez-Rodriguez, J., Kremling, A., and Gilles, E.D.: 'Dissecting the puzzle of life: Modularization of signal transduction networks', *Comput. Chem. Eng.*, in press
6 Saez-Rodriguez, J., Kremling, A., Conzelmann, H., Bettenbrock, K., and Gilles, E.D.: 'Modular analysis of signal transduction networks', *IEEE Contr. Syst. Mag.*, in press
7 Wolf, D.M., and Arkin, A.P.: 'Motifs, modules and games in bacteria', *Curr. Op. Microbiol.*, 2003, **6**, (2), pp. 125–134
8 Bisswanger, H.: 'Enzyme kinetics. Principles and methods' (Wiley-VCH, Weinheim, Germany, 2002)
9 Schneider, K.R., and Wilhelm, T.: 'Model reduction by extended quasi-steady-state approximation', *J. Math. Biol.*, 2000, **40**, (5), pp. 443–450
10 Roussel, M.R., and Fraser, S.J.: 'Invariant manifold methods for metabolic model reduction', *Chaos*, 2001, **11**, (1), pp. 196–206
11 Maeda, Y., Pakdaman, K., Nomura, T., Doi, S., and Sato, S.: 'Reduction of a model for an *Onchidium* pacemaker neuron', *Biol. Cybern.*, 1998, **78**, (4), pp. 265–276
12 Kremling, A., Fischer, S., Sauter, T., Bettenbrock, K., and Gilles, E.D.: 'Time hierarchies in the escherichia coli carbohydrate uptake and metabolism', *Biosystems*, 2004, **73**, (1), pp. 57–71
13 Falk, S., Guay, A., Chenu, C., Patil, S.D., and Berteloot, A.: 'Reduction of an eight-state mechanism of cotransport to a six-state model using a new computer program', *Biophys. J.*, 1998, **74**, (2 Pt 1), pp. 816–830
14 Obinata, G., and Andersion, B.D.O.: 'Model reduction for Control System Design' (Springer, London, 2001)
15 Scherpen, J.: 'Balancing for Nonlinear Systems'. PhD thesis, University of Twente, 1994
16 Romo, T.D., Clarage, J.B., Sorensen, D.C., and Phillips, G.N., Jr.: 'Automatic identification of discrete substates in proteins: singular value decomposition analysis of time-averaged crystallographic refinements', *Proteins*, 1995, **22**, (4), pp. 311–321
17 Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U.: 'Network motifs: Simple building blocks of complex networks', *Science*, 2002, **298**, pp. 824–828
18 Schweickhardt, T., and Allgöwer, F.: 'How nonlinear is nonlinear? an approach to nonlinearity quantification'. Proc. 7th Philips Conf. on Applications of Control Technology, 2003
19 Wiley, H.S., Shvartsman, S.Y., and Lauffenburger, D.A.: 'Computational modeling of the EGF-receptor system: a paradigm for systems biology', *Trends Cell. Biol.*, 2003, **13**, (1), pp. 43–50

20 Jorissen, R.N., Walker, F., Pouliot, N., Garrett, T.P., Ward, C.W., and Burgess, A.W.: 'Epidermal growth factor receptor: mechanisms of activation and signalling', *Exp. Cell. Res.*, 2003, **284**, (1), pp. 31–53
21 Ferrell, J.E., Jr.: 'Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs', *Trends Biochem. Sci.*, 1996, **21**, (12), pp. 460–466
22 Ferrell, J.E., Jr., and Xiong, W.: 'Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible', *Chaos*, 2001, **11**, pp. 227–236
23 Gong, Y., and Zhao, X.: 'Shc-dependent pathway is redundant but dominant in MAPK cascade activation by EGF receptors: a modeling inference', *FEBS Lett.*, 2003, **554**, (3), pp. 467–472

8 Appendix

8.1 Number of potential receptor species in the EGF network

First, we want to repeat the calculations introduced by Hlavacek *et al.* [2], additionally including the GAP binding and the internalisation. Therefore, the number of different states for each binding site of the receptor has to be determined. The EGF binding site can be bound to EGF or not, corresponding to two possible states. The GAP binding site can be bound to GAP or not, also corresponding to two possible states. The Grb2 binding site can be: (1) unphosphorylated, (2) phosphorylated, (3) bound to Grb2, (4) bound to Grb2 associated with Sos, corresponding to 4 different states. The Shc binding site can be: (1) unphosphorylated, (2) phosphorylated, (3) bound to Shc, (4) bound to phosphorylated Shc, (5) bound to Shc associated with Grb2, and (6) bound to Shc associated with Grb2 and Sos, corresponding to six different states. The number of monomeric species are therefore $2 \times 2 \times 4 \times 6 = 96$. The number of symmetric dimers is identical to the number of monomers (96), while the number of asymmetric dimers is $\binom{96}{2} = 4560$. Including receptor internalisation doubles the number of these species to 9504.

8.2 Number of all potential species in the whole EGF network

The number of species increases even more if we consider the whole EGF network as by Schoeberl *et al.* [3]. In this case the EGF binding site as well as the GAP binding site both have two possible states. The Grb2 binding site can be: (1) unphosphorylated, (2) phosphorylated, (3) bound to Grb2, (4) bound to Grb2 associated with Sos, (5) bound to Grb2 associated with Sos and Ras-GDP, and (6) bound to Grb2 associated with Sos and Ras-GTP, corresponding to six states. The Shc binding site can be (1) unphosphorylated, (2) phosphorylated, (3) bound to Shc, (4) bound to phosphorylated Shc, (5) bound to Shc associated with Grb2, (6) bound to Shc associated with Grb2 and Sos, (7) bound to Shc associated with Grb2, Sos and Ras-GDP, and (8) bound to Shc associated with Grb2, Sos and Ras-GTP, corresponding to eight states.

This leads to 192 monomers, 192 symmetric dimers and 18 336 asymmetric dimers. If we additionally include the MAP kinase cascade and receptor internalisation the number of possible species grows even more. Receptor internalisation nearly doubles the number of states, because almost all species exist in an internalised and non-internalised state. Only 15 species exist in one state (e.g. GAP, Grb2, Sos, Shc, Ras-GDP, Raf, MEK, ERK, Phosphatase1,…). All 18 720 receptor species and the remaining 17 species of the MAP kinase cascade exist in an internalised and non-internalised state. Therefore, the resulting number of species in the whole EGF signalling pathway is 37 489.

8.3 Analysis of pattern B

The linearised pattern B can be transformed into a transfer function of the form

$$G(s) = \frac{k(s+a_0)}{(s^2+bs+c)} = \frac{k(s+a_0)}{(s+a_1)(s+a_2)} \tag{29}$$

with

$$k = k_1 c_{A,s} \qquad a_0 = k_3 + k_{-2} \tag{30}$$

$$b = k_2 + k_{-1} + k_{-2} + c_{B,s}k_1 + k3 = b_0 + b_1 k_3 \tag{31}$$

$$\begin{aligned} c &= (k_2 + k_{-2})k_1 c_{B,s} + k_{-1}k_{-2} + (k_{-1} + k_2 + c_{B,s}k_1)k_3 \\ &= c_0 + c_1 k_3 \end{aligned} \tag{32}$$

For growing values of k_3 the limit for a_1 is

$$\begin{aligned} \lim_{k_3 \to \infty} \frac{1}{2}\left(b_0 + b_1 k_3 - \sqrt{(b_0 + b_1 k_3)^2 - 4(c_0 + c_1 k_3)}\right) \\ = \frac{c1}{b1} = k_{-1} + k_2 + c_{B,s}k_1 \end{aligned} \tag{33}$$

As written above, the condition $a_0 > a_1$ (see (27) and text below) implies the absence of a peak overshoot. $a_0 > a_1$ can be written for high values of k_3 according to (30) and (33) as

$$k_3 + k_{-2} > k_{-1} + k_2 + c_{B,s} \Longleftrightarrow \frac{k_3 + k_{-2}}{k_{-1} + k_2 + c_{B,s}} > 1 \tag{34}$$

Hence, for high enough values of k_3 the condition $a_0 > a_1$ will hold.