Published in IET Systems Biology
Received on 2nd March 2007
Revised on 19th July 2007
doi: 10.1049/iet-syb:20070012

ISSN 1751-8849

Multistability of signal transduction motifs

J. Saez-Rodriguez[1,4] *A. Hammerle-Fickinger*[2,4] *O. Dalal*[3] *S. Klamt*[4] *E.D. Gilles*[4] *C. Conradi*[4]

[1]*Harvard Medical School, Boston, MA 02115, U.S.A*
[2]*Centre of Life and Food Sciences Weihenstephan, Technical University of Munich, Germany*
[3]*Indian Institute of Technology, Bombay, Mumbai 400076, India*
[4]*Max-Planck-Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany*
E-mail: julio@hms.harvard.edu

Abstract: Protein domains are the basic units of signalling processes. The mechanisms they are involved in usually follow recurring patterns, such as phosphorylation/dephosphorylation cycles. A set of common motifs was defined and their dynamic models were analysed with respect to number and stability of steady states. In a first step, Feinberg's chemical reaction network theory was used to determine whether a motif can show multistationarity or not. The analysis revealed that, apart from double-step activation motifs including a distributive mechanism, only those motifs involving an autocatalytic reaction can show multistationarity. To further characterise these motifs, a large number of randomly chosen parameter sets leading to bistability was generated, followed by a bifurcation analysis of each parameter set and a statistical evaluation of the results. The statistical results can be used to explore robustness against noise, pointing to the observation that multistationarity at the single-motif level may not be a robust property; the range of protein concentrations compatible with multistationarity is fairly narrow. Furthermore, experimental evidence suggests that protein concentrations vary substantially between cells. Considering a motif designed to be a bistable switch, this implies that fluctuation of protein concentrations between cells would prevent a significant proportion of motifs from acting as a switch. The authors consider this to be a first step towards a catalogue of fully characterised signalling modules.

1 Introduction

Thanks to the spectacular development of molecular biology during the last years, our view of cellular processes in general, and signal transduction in particular, has evolved into a picture of fascinating, yet overwhelming, complexity, which suggests the use of mathematical analysis as a key tool to unravel it [1, 2].

Particularly, a modular approach, where networks are decomposed into subunits and subsequently analysed, has been proposed as a promising rationale for the analysis of large biochemical networks [3–6]. Since the concept of modularity can be applied to different levels of detail in a hierarchical manner [7], these subunits may comprise anything ranging from a single domain of a particular molecule to a whole organism. Usually, simple units are referred to as motifs, whereas larger components are named modules [8].

Even though the concept of modularity is widely accepted, a clear, unique criterion to identify modules is lacking [8]. Criteria may involve, among others, to be chemically isolated, active at a certain time scale or place, robust, statistically relevant [9] or clustered according to graph-theory methods [8, 10, 11]. Accordingly, there are a large number of works attempting to unravel the modularity of biochemical networks from all kinds of perspectives. Arguably, a reasonable criterion for decomposing signalling networks into modules from a system-theoretical point of view is to define units in a unidirectional manner, that is, without local bidirectional coupling between pairs of elements [7]. The resulting subsystems have properties which are independent of downstream elements and can be analysed using (systems) theoretical methods in a straightforward manner [6, 12].

The studies mentioned above follow a sort of 'top-down' approach: consider a protein interaction network and try to

unravel its modular properties. An inverse, 'bottom-up' perspective could also be revealing: consider the molecular architecture of the proteins involved in signalling systems with regard to the modularity. Recently, important modelling efforts have been made that consider protein domains as the building blocks of signalling processes (as is the case from a molecular perspective) [13–17]. To facilitate the understanding of signalling processes, their rigorous and detailed description is desirable. Unfortunately, rigour and detail go hand in hand with a combinatorial explosion of the number of states [12, 15]. However, recent work has demonstrated that this complexity can be significantly reduced without loss of exactitude if a domain-oriented approach is used. Moreover, applying this method, signal transduction networks show a modular structure where, if the domains do not interact allosterically, each module corresponds to a molecular domain and is connected in a unidirectional manner [12, 18].

Since the possible events taking place at a single domain are discrete, the number of possible motifs is limited. Therefore one can develop a construction kit of motifs, so that most signal transduction networks can be set up as an aggregation of those elements. Note that these motifs are minimal in the sense that they involve only one domain. Thus, they are different from those obtained from other approaches. For example, studies which uncover statistically relevant patterns provide different motifs, normally involving two to three domains [9].

Once such a kit is defined, it would be reasonable to analyse its elements systematically; upon thorough analysis of the units, the properties of combinations of them should be easier to understand. Stability would arguably be one of the first properties to examine, not only from a system-theoretical but also from a biological point of view, since it is related to important biological processes such as irreversible decision-making events controlling cell fate and differentiation processes [19, 20]. Once single motifs are well understood, the behaviour of combinations of them should be studied, to understand the emerging properties of large networks [21].

This is the central idea of this contribution, which is organised as follows.

We start with a brief review of the domain-oriented approach, and, in the light of this novel perspective, we define a construction kit of characteristic modules of signal transduction networks. The modules are implemented in ProMoT [22], a convenient tool since it allows one to set up models in a modular manner as well as to export them to the different analysis tools which will be used thereafter. The reactions are described in a mechanistic manner following the law of mass action. Besides, we assume throughout the paper that the total concentration of all the proteins is constant, that is, there is no degradation or transport. Subsequently, these motifs are systematically analysed with regard to their stability (Fig. 1).

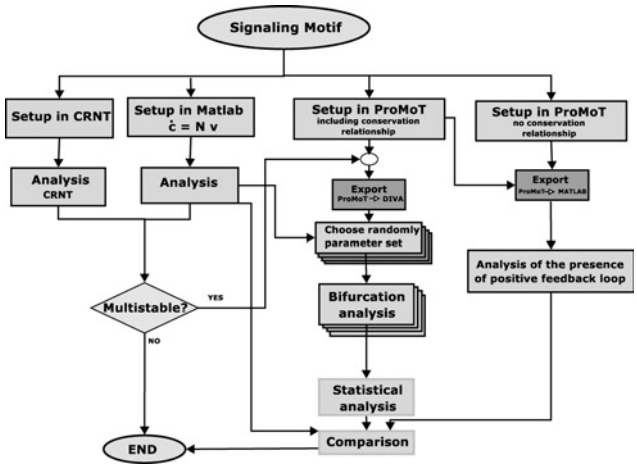

Figure 1 *Methodology followed in this study*

Once the modules are defined, the presence of positive feedbacks in the Jacobian of all motifs is determined
Parallelly, they are set up in CRNT Toolbox and ProMoT
Analysis with CRNT Toolbox indicates whether a motif can show multistationarity
In case a motif shows multistationarity, a large number of randomly generated parameters showing multistability is generated, and a detailed bifurcation analysis is performed for all of them in DIVA
Results are subsequently analysed statistically. Results of the different approaches are finally compared
See the main text and Methods section for details

First, the capability to show multistationarity is determined using Feinberg's chemical reaction network theory [23, 24] (CRNT). CRNT allows one to determine whether a certain (bio)chemical system can show multistationarity.

A key property of it is that it can determine the potential ability of a certain biochemical network to show multistationarity (for at least one set of parameter values) from the network structure, assuming that all kinetics are of the mass-action form. This makes it a useful tool when one faces systems where the structure, but not the kinetic parameters, is known [25, 26].

In case multistationarity is possible, the stability properties of the motif are explored using bifurcation analysis. Since we are interested in the general properties of these motifs and not only in the properties of a specific parameter-dependent realisation, a large number of parameter sets leading to multistationarity were determined. Subsequently bifurcation analysis was performed for each of these parameter sets. The resulting regions of multistationarity in parameter space were then statistically processed, both to assess the robustness of the multistable behaviour of a single module and to be able to compare robustness of different modules.

As a preliminary step, the existence of positive feedbacks (i.e. closed loops) in the incidence graph associated with the Jacobian matrix of every motif is explored (see Fig. 1 and supplementary information). The presence of positive feedback (PF) is a known requirement for multistability. Since

all motifs contain PF loops, multistability could not be excluded with this criterion in any motif.

2 Results

2.1 Domain-oriented approach underscores the importance of motifs in signalling networks

Signalling processes are mediated by receptors and adaptor proteins. A hallmark of these proteins is their ability to bind different molecules via different domains. The interface between proteins is built in a modular fashion, with few interactions between domains [27]. Therefore these domains can be considered as the fundamental elements of signal transduction [28].

If one considers all possible protein combinations (Fig. 2*a*), the number of feasible states explodes to thousands or even millions [12, 16, 18, 29]. This combinatorial complexity has typically been circumvented by assuming that only certain compounds are relevant (Fig. 2*b*). However, it is not possible to determine a priori which micro-states are the important ones [30], and therefore, a rigorous description has to include all possible states [15].

Recently, a new approach based on the macro-states (the states of the different domains) instead of the micro-states (the possible molecular combinations) has been proposed [16] which has been extended and formalised in [17]. The latter methodology provides a description equivalent to the microscopic one, but the resulting models are of a lower dimension and are structured in a modular manner, linking each module to a protein domain. Thus, the fact that protein domains are the actual units of signalling networks from a molecular perspective rationalises the set-up of mathematical models. Importantly, the

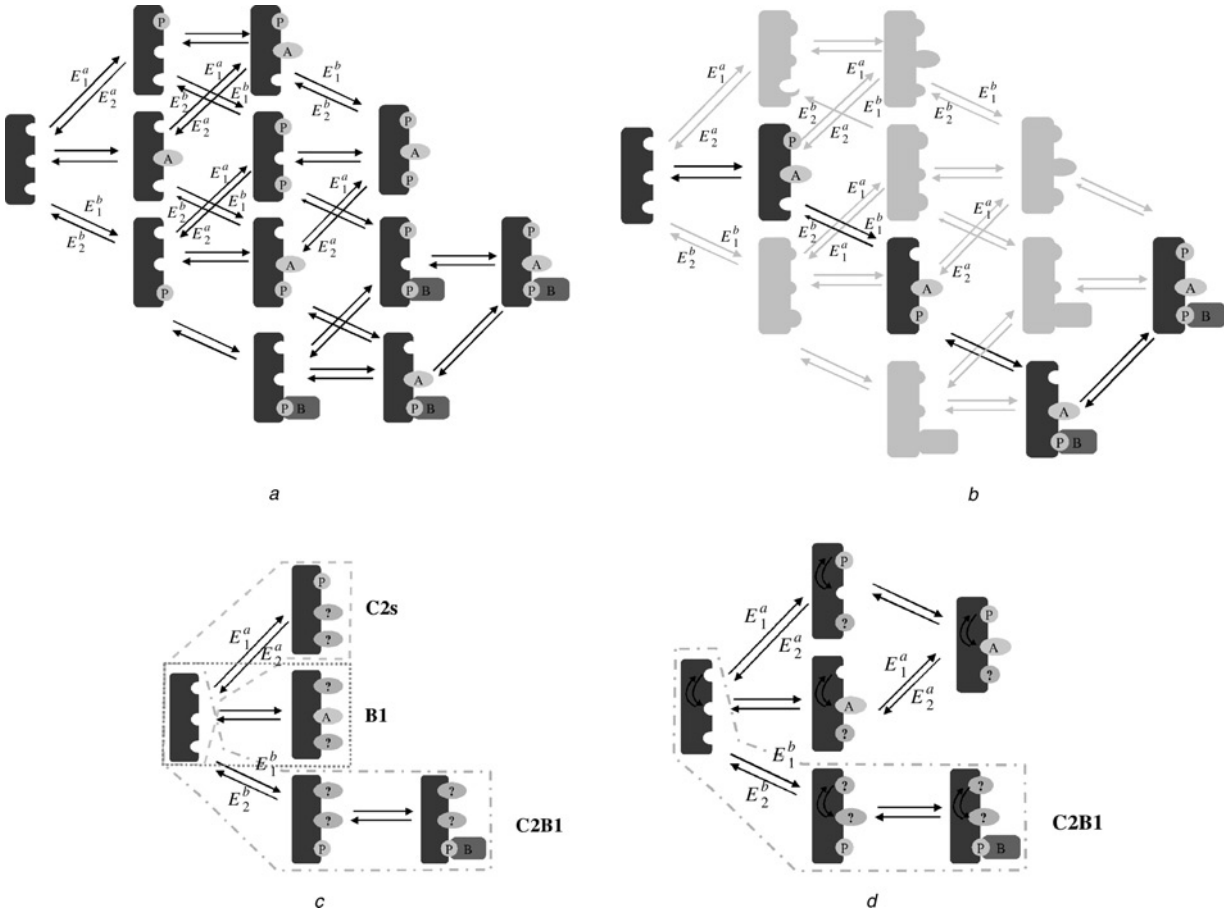

Figure 2 *Combinatorial complexity in signal transduction networks and their rigorous, modular description using a domain-oriented approach, illustrated by a protein with three domains*

First domain can be phosphorylated by an enzyme E_1
At the second domain, another protein *A* can bind and at the third, upon phosphorylation by an enzyme E_2, a protein *B* can bind
a Possible states (micro-states) are represented
b Traditionally, the modeller assumes an order of reactions which, however, is not always justified
Note that in this case, a separation into modules is not possible
c If the domains do not influence each other, they can be rigorously considered as independent entities (Figure 2(c)), corresponding to the simple motifs B1, C2s and C2B1 depicted in Figs. 3*d*, 3*a* and 3*e* respectively
d If two are somehow interconnected, they must be considered together, whereas the third can be considered alone

doi: 10.1049/iet-syb:20070012

concentrations of the individual complexes can be mapped back from the transformed states.

A key advantage is that the domains can be modelled independently, if they do not influence each other's long as the different domains in a molecule do not share reaction partners, the different domains in a molecule behave as independent entities. For example, in the system depicted in Fig. 2*a*, if the domains do not affect each other, the whole system can be replaced by the elementary motifs C2s (Fig. 3*a*), B1 (Fig. 3*d*) and C2B1(Fig. 3*e*), Fig. 2*c*.

However, if the state of one domain influences the properties of another (e.g. if binding of a protein to a certain domain changes the conformation of another one), they must be considered together (Fig. 2*d*). Note, however, that if, e.g. the phosphorylation of a domain converts this into an active catalytic centre which in turn can affect another one, but the latter not the former (e.g. in Fig. 2, if the phosphorylated form of the first domain is E_2), then the domains can be still considered independent.

Accordingly, to set up large signalling networks, one should wire together simple motifs belonging to a discrete set wherever domains can be considered independently, and only model them in a combined manner if an inter-domain influence exists.

In summary, a domain-oriented approach [17] provides a mathematically rigorous rationale for the biologically intuitive fact that protein domains are the building blocks of signalling networks.

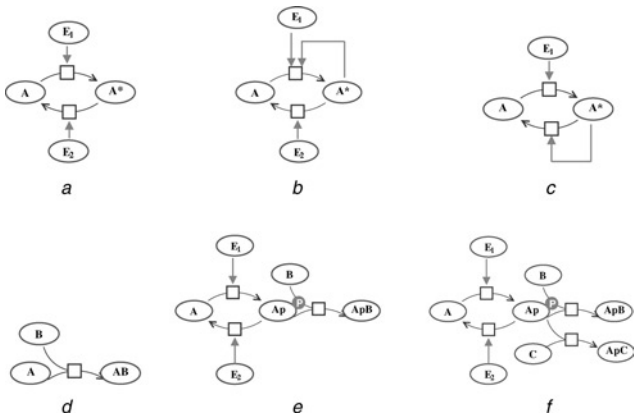

Figure 3 *Signal transduction motifs involving cycles of activation/deactivation and binding of domains*

a C2s: simple cycle
b C2p: cycle with PF
c C2n: cycle with negative feedback
d B1: binding of two domains
e C2B1: binding of one domain to an activated one
f C2B2: binding of two possible domains to an activated one

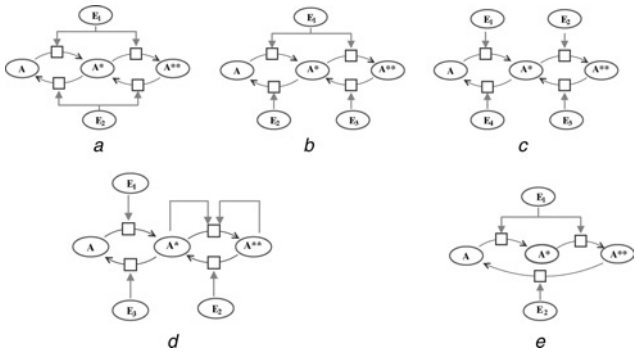

Figure 4 *Signal transduction motifs involving double cycles of activation/deactivation*

a C3dd: both reactions following a distributive mechanism
b C3di: one reaction follows a distributive mechanism and the other one is catalysed by two independent enzymes
c C3ii: all reactions are catalysed by independent enzymes
d C3sr: the second reaction is autocatalysed by its product
e C3dp: one reaction follows a distributive mechanism and the other one a processive one

2.2 Definition of a construction kit of signal transduction motifs

A construction kit of motifs composed of common recurring motifs was developed in ProMoT. ProMoT is an object-oriented modelling tool, which allows one to set up models in a hierarchical and modular manner [22]. Therefore based on this library of modules, large signalling networks can be modelled by simple drag-and-drop and subsequent wiring of the corresponding modules. Furthermore, ProMoT models can not only be exported using the Systems Biology Modelling Language [31], but also directly to Matlab and DIVA, where the analysis described below will be performed.

The elements of this construction kit (see Fig. 3, 4, and 5 and Table 1) describe recurring biochemical processes taking place at protein domains. The set represents the most common and relevant motifs. All reactions have been modelled following the mass-action law, and assuming no consumption or degradation of the molecules involved.

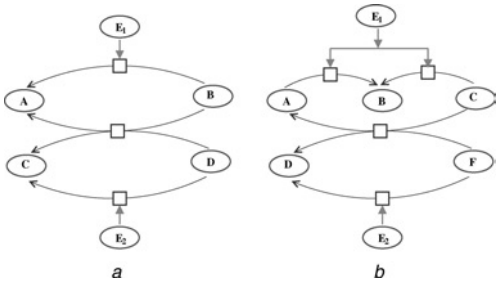

Figure 5 *Signal transduction motifs involving coupled cycles of activation/deactivation*

a C2C2: two single cycles coupled
b C3C2: one single and one double cycle coupled

Table 1 Textual definition of the motifs under study

C2s	*C2n*	*C2p*
$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$
$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$	$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$	$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$
	$A^* + A^* \underset{k_8}{\overset{k_7}{\rightleftharpoons}} A^*A^* \overset{k_9}{\rightarrow} A + A^*$	$A + A^* \underset{k_8}{\overset{k_7}{\rightleftharpoons}} AA^* \overset{k_9}{\rightarrow} A^* + A^*$
B1	*C2B1*	*C2B2*
	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$
$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AB$	$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$	$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$
	$A^* + B \underset{k_8}{\overset{k_7}{\rightleftharpoons}} A^*B$	$A^* + B \underset{k_8}{\overset{k_7}{\rightleftharpoons}} A^*B$
		$A^* + C \underset{k_{10}}{\overset{k_9}{\rightleftharpoons}} A^*C$
C3dd	*C3di*	*C3dp*
$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$
$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$	$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$	$A^{**} + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^{**}E_2 \overset{k_6}{\rightarrow} A + E_2$
$A^* + E_1 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} AE_1 \overset{k_9}{\rightarrow} A^{**} + E_1$	$A^* + E_1 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} AE_1 \overset{k_9}{\rightarrow} A^{**} + E_1$	$A^* + E_1 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} AE_4 \overset{k_9}{\rightarrow} A^{**} + E_1$
$A^{**} + E_2 \underset{k_{11}}{\overset{k_{10}}{\rightleftharpoons}} A^{**}E_2 \overset{k_{12}}{\rightarrow} A^* + E_2$	$A^{**} + E_3 \underset{k_{11}}{\overset{k_{10}}{\rightleftharpoons}} A^{**}E_3 \overset{k_{12}}{\rightarrow} A^* + E_3$	
C3ii	*C3sr*	*C3sr* (cont)
$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A^* + A^* \underset{k_8}{\overset{k_7}{\rightleftharpoons}} A^*A^* \overset{k_9}{\rightarrow} A^{**} + A^*$
$A^* + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_2 \overset{k_6}{\rightarrow} A + E_2$	$A^* + E_3 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*E_3 \overset{k_6}{\rightarrow} A + E_3$	$A^{**} + E_2 \underset{k_{11}}{\overset{k_{10}}{\rightleftharpoons}} A^{**}E_2 \overset{k_{12}}{\rightarrow} A^* + E_2$
$A^* + E_4 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} AE_4 \overset{k_9}{\rightarrow} A^{**} + E_4$	$A^* + A^{**} \underset{k_8}{\overset{k_7}{\rightleftharpoons}} A^*A^{**} \overset{k_9}{\rightarrow} A^{**} + A^{**}$	
$A^{**} + E_3 \underset{k_{11}}{\overset{k_{10}}{\rightleftharpoons}} A^{**}E_3 \overset{k_{12}}{\rightarrow} A^* + E_3$		
C2C2	*C3C2*	*C3C2* (cont)
$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A^* + E_1$	$A + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE_1 \overset{k_3}{\rightarrow} A* + E_1$	$B^* + E_2 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} B^*E_2 \overset{k_9}{\rightarrow} B + E_2$
$A^* + B \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^*B \overset{k_6}{\rightarrow} A + B*$	$A^{**} + B \underset{k_5}{\overset{k_4}{\rightleftharpoons}} A^{**}B \overset{k_6}{\rightarrow} A + B*$	$A^* + E_2 \underset{k_{11}}{\overset{k_{10}}{\rightleftharpoons}} A^*E_1 \overset{k_{12}}{\rightarrow} A^{**} + E_1$
$B^* + E_2 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} B^*E_2 \overset{k_9}{\rightarrow} B + E_2$		

Reactions define the modules depicted in Figs. 3–5

A central motif is a cycle of activation/deactivation [32], see Fig. 3. It typically involves a phosphorylation, which leads to the activation of the protein [33], and is counteracted by a dephosphorylation (see Fig. 3*a*). Another common realisation of this motif is via a guanine nucleotide exchange factor that promotes the exchange of guanosine diphosphate (GDP) by guanosine-triphosphate (GTP) and is opposed by a GTPase-activating protein.

doi: 10.1049/iet-syb:20070012

Sometimes, an activated motif can activate the same domain in another molecule (e.g. a kinase autophosphorylating itself in a regulatory side) leading to an auto-catalytic effect (Fig. 3*b*). On the other hand, the activated motif may be able to deactivate the same motif in another molecule, giving rise to a negative feedback (Fig. 3*c* Note that if an additional enzyme can also deactivate the motif, the system is symmetric, and thus regarding the purpose of this work equivalent, to the case C2p depicted in Fig. 3*b*). For example, a phosphatase (e.g. SHP1) which is activated by phosphorylation can in turn dephosphorylate (and thus deactivate) molecules of the same sort. Since these are cycles involving two states, we shall call C2 the generic case, and we add an additional letter (e.g. C2p) to specifically mention a variant of this motif.

In addition to activation/deactivation cycles, another essential motif of signal transduction is the binding of two domains (Fig. 3*d*), leading to the formation of a complex, as it is for example, the case of SH3 domains which bind to sequences of aminoacids (domains) rich in proline [33], and PH domains which bind to phospholipids [33] (In a sense, this could be interpreted as an activation/deactivation cycle, but it is reasonable to separate both cases here, since (i) no chemical change of the domain(s) takes place (the binding is a pure physical process) and (ii) for the activation/deactivation cycles only one domain is considered, while the other element (e.g. the phosphate group) is neglected.). This process can also take place combined with an activation/deactivation cycle (Fig. 3*e*), as is for example, the case of SH2 or PTB domains, which bind to certain phosphorylated motifs -so-called Tyrosine-based signalling motifs (TBSMs)–when they are phosphorylated [33]. Here, it may happen that two or more molecules compete to bind to the domain (see Fig. 3*f*).

It is common that the activation of a protein requires the phosphorylation of two binding sites. In principle, one would describe such a case as a combination of two domains.

However, in some cases it is known that the phosphorylation is not random and takes place in a succesive mannner. Therefore this case can be (and is virtually always) described as a lumped motif where a double activation cycle takes place. Since these motifs now involve three states connected by cycles of activation/deactivation, we shall call these modules C3. Often, both phosphorylation and dephosphorylation follow a distributive mechanism [26, 34], where the enzyme releases the substrate after the first binding (Fig. 4*a*).

In many cases, one enzyme is responsible for the activation of both domains, and one for the deactivation (Fig. 4*a*). However, in some cases, there is a specific enzyme involved in the activation and/or deactivation of each domain Figs. 4*b* and 4*c* (Note that the case of two enzymes involved in the activation and one in the deactivation is symmetric and thus, for the purpose of this work, equivalent to the case C3di depicted in Fig. 4*c*). Alternatively, one of both enzymes may follow a processive mechanism, where it may perform both modifications in one molecular encounter (Fig. 4*e*) [26, 34]. Furthermore, autocatalytic or autoinhibitory effects can also appear. We consider here one specific case, where the intermediate and fully activated form are both active and produce a PF via activation of the second reaction (Fig. 4*d*). This is the case, for example, of Src kinases, where the deactivated form (left) is phosphorylated at a negatively regulating site the fully activated is phosphorylated at a positive regulating site and the intermediate form not phosphorylated in any of them [35].

It is also common that cycles of activation/deactivation are coupled. An example widespread in bacteria is the two-component system, where one protein (the sensor) is phosphorylated upon perception of a stimulus and then the phosphate group is transferred to another protein (the receiver), resulting in its activation [36]. We consider two variants of this motif, one coupling two simple cycles (Fig. 5*a*), and one combining a double and a simple cycle (Fig. 5*b*).

2.3 CRNT analysis

The motifs described above (Figs. 3, 4 and 5) were then systematically analysed using CRNT. The results are displayed in Table 2. CRNT revealed that nine motifs (C2s, C2n, B1, C2B1, C2B2, C3ii, C3dp, C2C2,C3C2) cannot show multistationarity for any conceivable vector or rate constant. Except for the cases where a distributive mechanism is present (C3dd and C3di), only modules involving an explicit (i.e. 'visible' in the biochemical network) PF (produced by an autocatalytic reaction) are able to show multistationarity: motifs C2p (Fig. 3*b*) and C3sr (Fig. 4*d*). It should be noted that a larger set of variants of the motif C2p has been recently presented in [32]. Furthermore, even though the mere presence of a processive mechanism for either activation or deactivation excludes multistationarity (C3dp, Fig. 4*e*), multistationarity is possible if both steps are performed by independent enzymes (C3di, Fig. 4*b*).

The results for C3dd and C3dp were firstly presented in [34] (using bifurcation methods) and confirmed with CRNT in [26]. The requirement that a kinase mediates both steps of either the activation or deactivation was recently shown by analytical methods in [37]. Here, we extend these results in the sense that we show that the distributive activation is the only mechanism (of a set of alternatives including simple and double cycles, coupled cycles and binding of proteins) without an explicit autocatalytic feedback that has the ability to produce multistationarity. Furthermore, our analyses rely on a mass-action law expression for the enzymatic reactions, instead of the implicit quasi-steady state assumption of the Michaelis-Menten description used in the works above [34, 37]. This is an important generalisation as it is known that such assumptions can introduce important differences in the behaviour of the systems [38, 39].

Table 2 Summary of the results of the analysis of the motifs depicted in Figs. 3–4

Module	PF	δ	CRNT	$\bar{\Delta}_{A_T}$	$1 - F(-\bar{\Delta}_{A_T}/2, \bar{\Delta}_{A_T}/2)$	$\bar{\Delta}_{E_{1T}}$	$\bar{\Delta}_{E_{2T}}$
C2s	∃	1	—	—	—	—	—
C2p	∃	2	MSS	0.2708	0.0367–0.652	0.9037	0.3661
C2n	∃	1	—	—	—	—	—
B1	∃	0	—	—	—	—	—
C2B1	∃	1	—	—	—	—	—
C2B2	∃	1	—	—	—	—	—
C3dd	∃	2	MSS	0.2562	0.3831–0.6694	0.0808	0.0806
C3di	∃	2	MSS	0.0786	0.793–0.8958	0.039	0.0875
C3ii	∃	2	—	—	—	—	—
C3sr	∃	3	MSS	0.2773	0.3553–0.6440	–	0.3760
C3dp	∃	1	—	—	—	—	—
C2C2	∃	1	—	—	—	—	—
C3C2	∃	1	—	—	—	—	—

First, the existence of a positive feedback (PF) in the Jacobian was determined
Deficiency δ and the possibility of multistationarity was computed with the CRNT toolbox (see methods)
For the cases where multistationarity is possible (marked with MSS in the third column), a large ($\simeq$1000) set of parameters leading to multistability was computed, its area of bistability characterised with the parameters Δ_{A_T}, $\Delta_{E_{1T}}$ and $\Delta_{E_{2T}}$ (see (1)–(3), respectively) and the resulting values were statistically analysed
Statistical distributions are depicted in Figs. S1, S2 and S3 in the supplementary information
Here, the mean values $\bar{\Delta}_{A_T}$, $\bar{\Delta}_{E_{2T}}$, $\Delta_{E_{2T}}$ are presented
Additionally, the estimated fraction of cells which would lie out of the bistability region [$1 - F(-\bar{\Delta}_{A_T}, \bar{\Delta}_{A_T})$] (see main text and methods) is shown

Finally, a comparison of these results with those of the previous section suggests that, even though it is a requirement for multistability, the presence of PF is not a suitable criterion for identifying multistable (sub)systems, since it seems to be ubiquitous in signal transduction motifs.

2.4 Comparing robustness against concentration fluctuations by bifurcation analysis

It seems worthwhile to recall that CRNT can only provide information about *multistationarity* (i.e. the presence of more than one positive steady state for certain parameter values), but not about *multistability* (i.e. the presence of more than one stable steady state). Furthermore CRNT gives only information whether or not multistationarity is possible and, if possible, provides a pair of steady states and the corresponding parameter vector [40, 41]. In most cases, and certainly for the modules analysed herein, there will exist infinitely many parameter vectors such that multistationarity is possible.

Thus, for an in-depth exploration of multistationarity, a bifurcation analysis was performed for each network, where CRNT predicted multistationarity. To this end, continuation methods implemented in the software package DIVA [22] were used (Fig. 1). This numerical analysis revealed that each module, where CRNT predicted multistationarity, in fact admits three steady states over a wide range of parameter values. As, over the same parameter range, two of these steady states were stable, those modules are in fact bistable systems (in the appropriate parameter region). Note that, as this information was obtained by numerical methods, one cannot rule out other dynamical behaviour (e.g. limit cycles, different forms of multistationarity, etc.).

To describe the behaviour of a motif in general and not just for specific parameter values, a large set (around 1000) of randomly chosen parameters leading to multistationarity was generated (see methods and supplementary material), and subsequently evaluated statistically. Note that this approach, that is, the analysis of properties of signalling and gene networks using randomly generated parameters has been successfully applied in different contexts [42, 43, 44, 45, 46]. For each of these random parameter sets, bifurcation analysis revealed that the motifs are bistable and that the bifurcation is characterised by a limit point. Fig. 6 shows a characteristic example.

doi: 10.1049/iet-syb:20070012

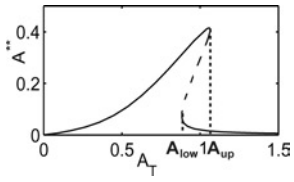

Figure 6 *Characteristic bifurcation analysis for the module C3di (Fig. 4b) with respect to A_T*

Lower (A_{low}) and up (A_{up}) bifurcation points, which are used to compute Δ_{A_T} (1) are depicted

In every case A_T, the total concentration of protein A was used as a bifurcation parameter. This parameter was chosen for the following biological reason. Since concentrations typically fluctuate, we argue that, in order to be a useful bistable switch, these motifs must keep bistability over a wide range of protein concentrations. Robustness against this sort of noise is postulated to be a hallmark of biological systems [47]. Under the following assumptions, the results of the bifurcation analysis were used to assess the robustness of each motif against the concentration fluctuations:

A1 Bistable motifs are 'biological switches'.

A2 Switches must be robust against concentration fluctuations in order to operate properly.

A3 For each motif, there exist multiple molecular realisations (domains from different proteins with the same structure). Each molecular realisation corresponds to a vector of rate constants (binding affinities, etc.). For each vector of rate constants leading to bistability, there exists a population of cells using this motif as a switching device.

Assumptions (A1) and (A2) imply a fourth assumption:

A4 All rate constants are in the 'bistability range', that is, they can be determined by the methods described in [48] and introduced in the supplementary material. Note that the resulting rate constants may not always be biochemically plausible.

In a population of cells, protein concentrations between cells fluctuate. Recent experimental results in living human cells revealed that the variability of protein concentrations between cells shows a standard deviation of about 15–30% [49]. Similar ranges have been also reported in yeast [50, 51], and bacteria [52]. This motivates the fifth assumption:

(A5) In a population of cells the protein concentration in individual cells is normally distributed with the mean close to the average of the protein concentrations taken over the population and a standard deviation of 15–30%.

Assumption (A2) implies the sixth assumption:

(A6) In every population of cells, realising a particular switching device, the average protein concentration is $c_b = A_{low} + A_{up}/2$, the centre of the bistability region (Fig. 6), thus guaranteeing maximal robustness.

Based on these assumptions, the following steps were taken to assess the robustness of multistationarity of those motifs where multistationarity is possible:

(i) For each motif, create a set of potential 'switching devices' by determining a random set of rate constants (i.e. molecular realisation) using the method described in [48] and in the supporting information. In these references, rate constants are parametrised by secondary parameters. These secondary parameters were normaly distributed with a mean of 0.693 and a standard deviation of 0.1. We graphically confirmed that the resulting rate constants were normally distributed as well.

(ii) For each element of the set (i.e. for each vector of parameters) determine the relative bistability range

$$\Delta_{A_T} = \frac{|A_{up} - A_{low}|}{((A_{low} + A_{up})/2)} \quad (1)$$

(iii) The relative bistability range is connected to the maximal change in protein concentration that the 'switching device' tolerates without losing the switching capability in the following way: if the protein concentration is expected to be close to the centre of the bistability region, the total protein concentration can vary at most by $\pm\Delta_{A_T}/2$ without losing the switching capability. Thus, the fraction of cells for which this particular 'switching device' is not operative was calculated as $1 - F(-\Delta_{A_T}/2, \Delta_{A_T}/2)$, where F is the normal cumulative distribution function for a normal distribution with a mean of 1 and a given standard deviation (in accordance with (A3), a standard deviation of 0.15–0.30 was used).

(iv) Finally, we computed the mean for all the vectors of parameters (molecular realisations) of a particular bistable motif. Table 2 contains the corresponding values.

Finally, under the assumptions (A1)–(A5), one obtains the following results regarding the robustness of multistationarity:

A. All modules show an average value of around 25% for $\bar{\Delta}_{A_T}$, except C3di, which shows a much lower value ($\simeq$8%). Note that in case of C3di (Fig. 4*b*), a single distributive double-step activation mechanism is responsible for multistationarity. In the other cases, either two double-step activation mechanisms (C3dd, Fig. 4*a*) or an explicit autocatalytic activation (C2d and C3sr, Figs. 3*b* and 4*d*) are responsible for multistationarity. These results suggest that an autocatalytic activation is a stronger (i.e. more robust to concentration fluctuations) mechanism to cause multistationarity than a distributive activation, since one of

autocatalytic mechanisms leads to similar ranges of Ram as two distributive mechanisms. (Note that one autocatalytic mechanism leads to similar ranges of $\bar{\Delta}_{\mathcal{A}_T}$ as two distributive mechanisms.)

B. The results summarised in Table 2 indicate that, under the assumption (A3), for an implementation of a switch with relative concentration range $\bar{\Delta}_{\mathcal{A}_T}$ a significant fraction of cells can be expected to have a value of $\mathcal{A}_T$ outside $\bar{\Delta}_{\mathcal{A}_T}$; that is, a significant fraction of cells would not be able to operate in the bistable modus, for any of the four motifs able to be bistable (see Table 2, column six).

C. If one considers the most robust implementations of each motif, (i.e. those with the highest values of $\Delta_{\mathcal{A}_T}$, see Figure S1 in the supplementary information) one sees that there are cases with a remarkably high $\Delta_{\mathcal{A}_T}$, particularly for C3dd ($\cong$0.95). Thus, such a realisation, if it was feasible from a physical point of view, would be quite a robust one, since only for $\simeq$2%–11% of the cells the value of $\mathcal{A}_T$ would lie out of the bistability region.

It should be noted that these values correspond to protein amounts which do not necessarily correlate to concentration because of volume changes, and that we assume a normal distribution which is probably not the case. Therefore these results should be considered as rough estimations. In summary, the use of single-level bistable switches does not seem to be a robust design, as there would be a significant part of the cells where it would not be operative, even though particular implementations may be done.

Bifurcation analysis was also performed with respect to enzymes E_1 (activating) and E_2 (deactivating) leading to relative concentration ranges $\Delta_{E_{1T}}$ and $\Delta_{E_{2T}}$, defined similar to (1)

$$\Delta_{E_{1T}} = \frac{|E_{1,\text{low}} - E_{1,\text{up}}|}{((E_{1,\text{low}} + E_{1,\text{up}})/2)} \tag{2}$$

and

$$\Delta_{E_{2T}} = \frac{|E_{2,\text{low}} - E_{2,\text{up}}|}{((E_{2,\text{low}} + E_{2,\text{up}})/2)} \tag{3}$$

Here, however, the interpretation is different. Rather than robustness against fluctuations in the concentration of the corresponding protein, $\Delta_{E_{1T}}$ and $\Delta_{E_{2T}}$ provide information about the range of operativity of the switch; they indicate for which input values (E_1 and E_2) the module shows bistability. In this context, the range of multistationarity is again much narrower for modules where multistationarity relies on a distributive mechanism, see Table 2. Note that, for C3sr, the parameter Δ E_1 cannot be computed; for an input higher than a certain critical value, the system is 'shifted' to the right cycle and it converts into a single cycle without input and an autocatalytic reaction. Thus, the system is, autonomous and does not depend on the input E_1 (see Figure S4 in the supplementary information).

3 Conclusions

In this contribution, a set of signal transduction motifs has been analysed with respect to the ability to show multistationarity, a non-linear phenomenon involved in key biological processes, particularly in the handling of signals determining cellular decisions. Therefore the identification of biochemical motifs related to multistationarity can provide important insights into the rationale behind signal processing and the composition of large signal transduction networks from these small motifs.

We have analysed a specific set of motifs present in relevant signalling networks, such as those processing growth factors and cytokines. According to a new domain-oriented modelling strategy [17], these motifs can be combined to set up a variety of signalling networks.

Our goal was to analyse multistability under conditions as realistic and detailed as possible. Thus, we have used a mass-action-law description for the enzymatic reaction. Thereby, we avoid the commonly not fulfilled and dangerous assumptions implicit in reduced expressions like Michaelis–Menten kinetics [38, 39], which have been used in other related works [34, 37].

For the analysis, a number of methods (bifurcation analysis, identification of PF loops, Feinberg's CRNT and a novel method to characterise the parameter space where multistability is possible [48]) have been applied, and their results have been compared. Our work shows the applicability of CRNT to signal transduction systems; it provides a fast and reliable method to uncover mechanisms leading to multistationarity (limited by the number of compounds that the toolbox can handle). To be more precise about the nature of the multistationarity, bifurcation methods are required; in our case, they revealed that bistability (a particular form of multistationarity of high biological relevance) was present in all motifs showing multistationarity. Furthermore, the statistical analysis of the bifurcations of these motifs for different parameter sets illustrates the new possibilities which opens a method as that described in [48], which allows one to analytically characterise the parameter space where multistationarity takes place. Although CRNT provides one set of parameters where multistability takes place, only with the method of [48] it is possible to obtain a sufficiently large number of different vectors of rate constants to perform the quantification of robustness.

On the other hand, the presence of PF in the incidence matrix associated with the Jacobian matrix, even though being a condition for multistationarity, does not seem to be a conclusive tool since PF loops are omnipresent in the Jacobian of virtually any signal transduction network. A more sophisticated analysis of the nature of the feedback loops, however, could help to determine which positive feedbacks have the potential to induce multistability [53].

doi: 10.1049/iet-syb:20070012

Our analysis shows under which conditions multistationarity can be expected and suggests that, in general, it may not be very robust against concentration fluctuations. Since the noisy nature of signalling and gene expression processes is a well-established fact [54, 55], the results may pose a word of caution with regard to the biological relevance of single-motif bistability, particularly in the light of recent experimental data in human cells [49], yeast [50] and bacteria [52]; the cell-to-cell variation in the protein level within a cell population would shift many cells out of the bistability range of a certain switch.

It may be, however, that single-motif switches, when embedded in cascades with PF regulation, expand their parameter space of bistability [34]. Alternatively, such a switch would be plausible if the concentration of the corresponding protein were subjected to strict regulation. Additionally, it would be interesting to know to what extent the parameter sets leading to a large Δ_{A_T} are physiologically plausible; if they were, it would be tempting to speculate whether they are *de facto* implemented in cellular networks.

Our analysis also suggests that autocatalytic reactions are stronger promoters of multistationarity than the competition between substrates and enzymes in distributive double-step activation processes. This would argue for a more common use of the former mechanism to obtain robust bistable behaviour.

This contribution provides a first step towards a characterisation of signal transduction building blocks; we have analysed only one (yet essential) property, but certainly other properties such as monotony, dynamics and input/output behaviour should be thoroughly studied. To this end, some previous works exist, mainly applied to activation/deactivation cycles [6, 42, 56, 57, 58].

One field where these results can be of interest is that of synthetic biology: while designing biochemical systems that are able to perform a particular task, very often bistable behaviour is required, for example to include an irreversible decision-making step. The results of this study may help to find suitable motifs to achieve this property.

Another direction of our future work regards the analysis of multistability and other properties of aggregation of motifs. It is well known that coupling of modules can lead to the appearance of new, emergent properties [21]. Therefore probably unexpected and exciting discoveries wait for us along this path.

4 Materials and methods

A number of tools have been used and compared in this contribution (see Fig. 1 for a description of their interconnections). First, the methods will be succinctly listed. Later on the method used to obtain the parameter sets leading to multistability will be explained in more detail.

- To compute the feedback loops, for each motif we first computed the Jacobian matrix $\boldsymbol{J} \in \mathbb{R}^{n\times n}$ with

$$\boldsymbol{J}_{ij} = \frac{\partial f_i}{\partial c_j}$$

associated with the corresponding system of differential equations

$$\dot{x} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = f(\vec{x}, \vec{u}, \vec{p})$$

in Matlab. The algorithm for the detection of feedbacks in the Jacobian matrix is embedded in CellNetAnalyzer [59], which can be downloaded from [60].

- Feinberg's CRNT was applied using CRNT Toolbox 1.1 (see [61] to download the software). For an introduction to Chemical Reaction Network from the perspective of systems biology, we refer the reader to [26], and for a deeper understanding of the theory itself to [23, 24].

- The differential equation models were set up with ProMoT [22], which can be obtained from [62].

- Bifurcation analysis was performed using continuation methods implemented in the software tool DIVA [22]. Although DIVA relies in part on commercial software, Diana (the successor of DIVA), can be obtained for free from [62].

- To estimate the proportion of cells outside the bistable region, we argued, as described in the main text, that if a certain motif has been optimised through evolution to be a switch, it should be as robust as possible against fluctuations of total concentration A_T, and thus have a value of A_T close to the centre of the bistability region. Therefore the proportion of cells with a value of A_T outside the bistability region would be those with

$$|A_T - \frac{A_{low} + A_{up}}{2}| > \frac{\Delta A_T}{2} \quad (4)$$

Assuming (as a rough approximation) a normal distribution with a mean of 0 and the typical standard deviation of the protein concentration, one can compute the amount of cells fulfilling (4) as $1 - F(-\bar{\Delta}_{A_T}/2, \bar{\Delta}_{A_T}/2)$, where F is the normal cumulative distribution function for a normal distribution with a mean of 0 and a given standard deviation (Fig. 7). Since it is the closest information available, we assumed for estimation purposes the standard deviation of the concentration to be similar to that of the total amount of protein (0.15–0.30 for mammal cells [49]). It should be noted that these values correspond to protein amounts which do not necessarily correlate to concentration because of volume changes, and that we assume a normal distribution which is probably not the case. Therefore these

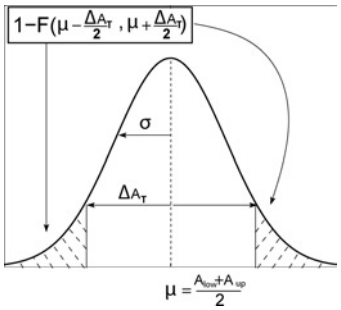

Figure 7 *Schematic representation of the procedure used to estimate the amount of cells out of the bistability regime*

Assuming a normal distribution with mean $\mu = A_T^1 + A_T^2/2$, for a certain standard deviation σ the amount of cells can be computed as $1 - F(\mu - \bar{\Delta}_{A_T}/2,\ \mu + \bar{\Delta}_{A_T}/2)$

results should be considered as a rough estimation. These statistical evaluations of the results were performed with Matlab.

4.1 Parameter sets leading to multistationarity

In [48] a model for the activation an MAPK(K) is analysed with respect to its ability to admit multiple steady states. For this particular model, it is possible to derive a parametrisation of all rate constants k_i that are associated with the existence of at least two positive steady states. The same method has been applied here to those motifs where CRNT predicts multistationarity to obtain such a parametrisation of parameter values. The supplementary material for this paper contains this parametrisation for every motif together with a brief description of how it was obtained.

For completeness, this procedure is briefly described. To avoid lengthy mathematical expressions and thus enhance readability and highlight systems-biological consequences of these equations, only a small and simple reaction network is considered. However, this method can be applied to all networks discussed in this paper. For a detailed discussion of the method we refer to [48] and the accompanying supplementary material.

Consider the following reactions between two species A and B

$$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2A \tag{5}$$

This network is a Deficiency Zero network in the terminology of Feinberg's CRNT and, thus, multistationarity is excluded for the network [40]. Nonetheless it is well suited to demonstrate the methodology used in [48], as all elements of this methodology can be applied to this network. To apply the procedure, consider the system of ODEs defined by this network in the following notation (where x_1 is the concentration of A and x_2 the concentration of B)

$$\dot{\boldsymbol{x}} = \boldsymbol{N}\boldsymbol{v}(\boldsymbol{k}, \boldsymbol{x}) \tag{6a}$$

where

$$\boldsymbol{N} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \tag{6b}$$

$$\boldsymbol{v}(\boldsymbol{k}, \boldsymbol{x}) = (k_1 x_1 x_2,\ k_2 x_1^2)^{\mathrm{T}} \tag{6c}$$

and the conservation relation

$$x_1 + x_2 = \text{const.} \tag{6d}$$

Multistationarity is given, if two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ with $\boldsymbol{N}\boldsymbol{v}(\boldsymbol{k},\ \boldsymbol{a}) = 0$, $\boldsymbol{N}\boldsymbol{v}(\boldsymbol{k},\ \boldsymbol{b}) = 0$ and $a_1 + a_2 = b_1 + b_2$ exist (i.e. $\boldsymbol{a}$ and $\boldsymbol{b}$ satisfy the conservation relation (6d) for the same right hand side). As we are only interested in positive steady states (i.e. $\boldsymbol{a},\ \boldsymbol{b} \in \mathbb{R}^2_{>0}$), the steady-state equations are equivalent to $\boldsymbol{v}(\boldsymbol{k},\ \boldsymbol{a}) = \boldsymbol{E}\ \lambda$ and $\boldsymbol{v}(\boldsymbol{k},\ \boldsymbol{b}) = \boldsymbol{E}\ \nu$, where $\boldsymbol{E} = (1,\ 1)^{\mathrm{T}}$ is a generator of the pointed polyhedral cone $\ker(\boldsymbol{N}) \cap \mathbb{R}^2_{>0}$ (note that for this network a generator of $\ker(\boldsymbol{N}) \cap \mathbb{R}^2_{>0}$ is also a basis for $\ker(\boldsymbol{N})$, in general this need not be the case).

This leads to the following equations

$$k_1 a_1 a_2 = \lambda \quad k_2 a_1^2 = \lambda \tag{7a}$$

$$k_1 b_1 b_2 = \nu \quad k_2 b_1^2 = \nu \tag{7b}$$

The next step is to subtract the equations in a_i from the corresponding equations in b_i and applying the natural logarithm. If the definition $\mu_i := \ln b_i/a_i$ is used, the result is a system of equations in μ_i and λ, ν

$$\mu_1 + \mu_2 = \ln\frac{\nu}{\lambda} \quad 2\mu_1 = \ln\frac{\nu}{\lambda} \tag{8}$$

The solution to these equations is given by

$$\boldsymbol{\mu} = \frac{1}{2}\ln\frac{\nu}{\lambda}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{9}$$

with ν, $\lambda > 0$. Thus, $\ln \nu/\lambda \in \mathbb{R}$ and $\kappa := 1/2 \ln \nu/\lambda$ can be interpreted as a free parameter. Therefore $\boldsymbol{\mu} = \kappa\ \boldsymbol{M}$, $\boldsymbol{M} = (1,\ 1)^{\mathrm{T}}$ defines a linear subspace of $\mathbb{R}^2$. Any element of this subspace gives rise to two positive vectors $\boldsymbol{a}$ and $\boldsymbol{b}$: choose $\kappa \in \mathbb{R}$ and $\boldsymbol{a} \in \mathbb{R}^2_{>0}$ and define $\boldsymbol{b} = e^{\kappa}(a_1,\ a_2)^{\mathrm{T}}$. Then $\boldsymbol{a}$ and $\boldsymbol{b}$ are steady states of (6a), if the following vector of rate constants is used (obtained by solving, for e.g. (7a) for $(k_1,\ k_2)^{\mathrm{T}}$)

$$(k_1,\ k_2)^{\mathrm{T}} = \left(\frac{\lambda}{a_1 a_2},\ \frac{\lambda}{a_1^2}\right)^{\mathrm{T}} \tag{10}$$

Note that, as λ is free, each pair $\boldsymbol{a}$, $\boldsymbol{b}$ is associated with an infinite set of parameter values.

doi: 10.1049/iet-syb:20070012

Even though (9) defines an infinite set of pairs $\boldsymbol{a}$, $\boldsymbol{b}$ (as (9) defines a linear subspace) the system (6a), (6d) does not admit multiple steady states: there exists no pair $\boldsymbol{a}$, $\boldsymbol{b}$ that satisfies (6d) for the same right hand side. To see this, consider $b_1 + b_2 = e^{\kappa}(a_1 + a_2) \neq a_1 + a_2$ (unless, of course, $\kappa = 0$; however $\boldsymbol{a} = \boldsymbol{b}$ in this case).

If one randomly generates 1000 different values for κ, then one obtains 1000 different pairs of steady states $(\boldsymbol{a}, \boldsymbol{b})_i$, $i = 1, \ldots, 1000$ and, for each of these pairs, (10) defines a vector of rate constants $\boldsymbol{k} = (k_1, k_2)^{\mathrm{T}}$, such that $(\boldsymbol{a}, \boldsymbol{b})_i$ are steady states (note that one still can choose the parameter λ in (10)). Thus, one obtains triplets $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{k}_i, i = 1, \ldots, 1000)$, each containging a pair of steady states and the corresponding vector of rate constants for network (5).

CRNT predicts multistationarity for the networks given in Figs. 3*b*, 4*a*, 4*b* and 4*d*. All of those networks were treated similar to network (5) (the corresponding equations are given in the supplementary material to this article):

1. For each of the networks, a parametrisation of steady-state pairs was obtained by applying the steps described above.

2. For each of these parametrisations a set of 1000 steady-state pairs was created randomly: $(\boldsymbol{a}, \boldsymbol{b})_i^{(j)}$, $i = 1, \ldots, 1000$ and $j = 1, \ldots, 4$, where $j = 1$ identifies the network given in Fig. 3*b*, $j = 2$ the network given in Fig. 4*a*, $j = 3$ the network given in Fig. 4*b* and $j = 4$ the network given in Fig. 4*d*.

3. Using parametrisations of the rate constants (like (10) for network (5)), a vector of rate constants was obtained for each pair of steady states. Thus, one obtains triplets $\tau_i^{(j)} := (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{k})_i^{(j)}$, $i = 1, \ldots, 1000, j = 1, \ldots, 4$, for which networks in Figs. 3*b*, 4*a*, 4*b* and 4*d* show multistationarity.

4. For every network and every triplet $\tau_i^{(j)}$, a numerical continuation of steady states was performed, using A_{T} as the bifurcation parameter. Bistability was confirmed numerically for each $\tau_i^{(j)}$.

5. Lower and upper turning points $A_{\mathrm{low},i}^{(j)}$ and $A_{\mathrm{up},i}^{(j)}$ of the numerical continuation were determined (see Fig. 6, where turning points are labelled LP, limit point). Using $A_{\mathrm{low},i}^{(j)}$ and $A_{\mathrm{up},i}^{(j)}$, for each $\tau_i^{(j)}$, the relative range of bistability was determined as

$$\Delta_{A_{\mathrm{T}},i}^{(j)} = \frac{|A_{\mathrm{up},i}^{(j)} - A_{\mathrm{low},i}^{(j)}|}{(A_{\mathrm{low},i}^{(j)} + A_{\mathrm{up},i}^{(j)})/2}$$

6. Finally $\bar{\Delta}_{A_T}^{(j)}$, the mean of all $\Delta_{A_{T,i}}^{(j)}$ was calculated for each network.

The steady-state equations were solved manually, and the Lemma 1 given in the supplementary information of an algorithm was implemented in Matlab (available upon request).

5 Acknowledgments

JSR, SK and EDG acknowledge the support from German Ministry of Research and Education (Hepatosys), the German Research Society (FOR521), and the Ministry of Education of Saxony-Anhalt (Research Focus Dynamical Systems). JSR thanks Martin Ginkel for help on the set-up of the ProMoT models.

Competing interests. The authors have declared that no competing interests exist.

6 References

[1] PAPIN J.A., HUNTER T., PALSSON B.O., SUBRAMANIAM S.: 'Reconstruction of cellular signalling networks and analysis of their properties', *Nat. Rev. Mol. Cell. Biol.*, 2005, **6**, (2), pp. 99–111

[2] SONTAG E.: 'Molecular Systems Biology and Control', *Euro. J. Control*, 2005, 11, pp. 396–435

[3] HARTWELL L.H., HOPFIELD J.J., LEIBLER S., MURRAY A.W.: 'From molecular to modular cell biology', *Nature*, 1999, **402**, (6761), pp. C47–C52

[4] LAUFFENBURGER D.A.: 'Cell signaling pathways as control modules: complexity for simplicity?', *Proc. Natl. Acad. Sci. USA*, 2000, **97**, (10), pp. 5031–5033

[5] PAPIN J.A., REED J.L., PALSSON B.O.: 'Hierarchical thinking in network biology: the unbiased modularization of biochemical networks', *Trends Biochem. Sci.*, 2004, **29**, (12), pp. 641–647

[6] SAEZ-RODRIGUEZ J., KREMLING A., CONZELMANN H., BETTENBROCK K., GILLES E.D.: 'Modular analysis of signal transduction networks', *IEEE Contral. Syst. Mag.*, 2004, **24**, (4), pp. 35–52

[7] SAEZ-RODRIGUEZ J., KREMLING A., GILLES E.D.: 'Dissecting the puzzle of life: modularization of signal transduction networks', *Comput. Chem. Eng.*, 2005, **29**, (3), pp. 619–629

[8] WOLF D.M., ARKIN A.P.: 'Motifs, modules and games in bacteria', *Curr. Opin. Microbiol.*, 2003, **6**, (2), pp. 125–134

[9] MILO R., SHEN-ORR S., ITZKOVITZ S., KASHTAN N., CHKLOVSKII D., ALON U.: 'Network motifs: simple building blocks of complex networks', *Science*, 2002, 298, pp. 824–828

[10] BARABASI A.L., OLTVAI Z.N.: 'Network biology: understanding the cell's functional organization', *Nat. Rev. Genet.*, 2004, **5**, (2), pp. 101–113

[11] NEWMAN M.E.J.: 'Modularity and community structure in networks', *Proc. Natl. Acad. Sci. USA*, 2006, **103**, (23), pp. 8577–8582

[12] CONZELMANN H., SAEZ-RODRIGUEZ J., SAUTER T., BULLINGER E., ALLGÖWER F., GILLES E.D.: 'Reduction of mathematical models of signal transduction networks: simulation-based approach applied to EGF receptor signaling', *Syst. Biology*, 2004, **1**, (1), pp. 159–169

[13] LOK L., BRENT R.: 'Automatic generation of cellular reaction networks with moleculizer 1.0', *Nat. Biotechnol.*, 2005, **23**, (1):131–136)

[14] SHIMIZU T.S., LE NOVERE N., LEVIN M.D., BEAVIL A.J., SUTTON B.J., BRAY D.: 'Molecular model of a lattice of signalling proteins involved in bacterial chemotaxis', *Nat. Cell. Biol.*, 2000, **2**, (11), pp. 792–796

[15] BLINOV M.L., FAEDER J.R., GOLDSTEIN B., HLAVACEK W.S.: 'BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains', *Bioinformatics*, 2004, **20**, (17), pp. 3289–3291

[16] BORISOV N.M., MARKEVICH N.I., HOEK J.B., KHOLODENKO B.N.: 'Signaling through receptors and scaffolds: independent interactions reduce combinatorial complexity', *Biophys. J.*, 2005, **89**, (2), pp. 951–966

[17] CONZELMANN H., SAEZ-RODRIGUEZ J., SAUTER T., KHOLODENKO B.N., GILLES E.D.: 'A domain-oriented approach to the reduction of combinatorial complexity in signal transduction networks', *BMC Bioinformatics*, 2006, 7, p. 34

[18] SAEZ-RODRIGUEZ J., CONZELMANN H., SAUTER T., KHOLODENKO B.N., GILLES E.D.: 'Domain-oriented and modular approaches to the reduction of mathematical models of signaling networks', in KUMMER U. PAHLE J. SUROVTSOVA I. ZOBELEY J. (EDs) '4th workshop on computation of biochemical pathways and genetic networks', (Logos-Verlag, Berlin, 2005), pp. 13–20

[19] EISSING T., CONZELMANN H., GILLES E.D., ALLGOWER F., BULLINGER E., SCHEURICH P.: 'Bistability analyses of a caspase activation model for receptor-induced apoptosis', *J. Biol. Chem.*, 2004, **279**, (35), pp. 36892–36897

[20] XIONG W., FERRELL J.E.J.: 'A positive-feedback-based bistable 'memory module' that governs a cell fate decision', *Nature*, 2003, **426**, (6965), pp. 460–465

[21] BHALLA U.S., IYENGAR R.: 'Emergent properties of networks of biological signaling pathways', *Science*, 1999, **283**, (5400), pp. 381–387

[22] GINKEL M., KREMLING A., NUTSCH T., REHNER R., GILLES E.D.: 'Modular modeling of cellular systems with ProMoT/Diva', *Bioinformatics*, 2003, **19**, (9), pp. 1169–1176

[23] FEINBERG M.: 'The existence and uniqueness of steady states for a class of chemical reaction networks', *Arch. Ration. Mech. Anal.*, 1995, **132**, (4), pp. 311–370

[24] FEINBERG M.: 'Multiple steady states for chemical reaction networks of defi-ciency one', *Arch. Ration. Mech. Anal.*, 1995, **132**, (4), pp. 371–406

[25] BAILEY J.E.: 'Complex biology with no parameters', *Nat. Biotechnol*, 2001, **19**, (6), pp. 503–504

[26] CONRADI C., SAEZ-RODRIGUEZ J., GILLES E.D., RAISCH J.: 'Using chemical reaction network theory to discard a kinetic mechanism hypothesis', *IEE Proc Syst. Biology*, 2005, **152**, (4), pp. 243–248

[27] REICHMANN D., RAHAT O., ALBECK S., MEGED R., DYM O., SCHREIBER G.: 'The modular architecture of protein-protein binding interfaces', *Proc. Natl. Acad. Sci., USA*, 2005, **102**, (1), pp. 57–62

[28] PAWSON T., NASH P.: 'Assembly of cell regulatory systems through protein interaction domains', *Science*, 2003, **300**, (5618), pp. 445–452

[29] HLAVACEK W.S., FAEDER J.R., BLINOV M.L., PERELSON A.S., GOLDSTEIN B.: 'The complexity of complexes in signal transduction', *Biotechnol. Bioeng.*, 2004, **84**, (7), pp. 783–794

[30] FAEDER J.R., BLINOV M.L., GOLDSTEIN B., HLAVACEK W.S.: 'Combinatorial complexity and dynamical restriction of network flows in signal transduction', *Syst. Biology*, 2005, **2**, (1), pp. 4–15

[31] HUCKA M., FINNEY A., SAURO H.M., BOLOURI H., DOYLE J.C., KITANO H.: 'The systems biology markup language, (SBML): a medium for representation and exchange of biochemical network models', *Bioinformatics*, 2003, **19**, (4), pp. 524–531

[32] KHOLODENKO B.N.: 'Cell-signalling dynamics in time and space', *Nat. Rev. Mol. Cell. Biol.*, 2006, **7**, (3), pp. 165–176

[33] KRAUSS G.: 'Biochemistry of signal transduction and regulation' (Wiley-VCH, 2003)

[34] MARKEVICH N.I., HOEK J.B., KHOLODENKO B.N.: 'Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades', *J Cell. Biol.*, 2004, **164**, (3), pp. 353–359

[35] PALACIOS E.H., WEISS A.: 'Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation', *Oncogene*, 2004, **23**, (48), pp. 7990–8000

[36] STOCK A.M., ROBINSON V.L., GOUDREAU P.N.: 'Two-component signal transduction', *Annu. Rev. Biochem.*, 2000, 69, pp. 183–215

doi: 10.1049/iet-syb:20070012

[37] ORTEGA F., GARCS J.L., MAS F., KHOLODENKO B.N., CASCANTE M.: 'Bistability from double phosphorylation in signal transduction. Kinetic and structural requirements', *FEBS J.*, 2006, **273**, (17), pp. 3915–3926

[38] MILLAT T., BULLINGER E., ROHWER J., WOLKENHAUER O.: 'Approximations and their consequences for dynamic modelling of signal transduction pathways', *Math. Biosci.*, 2006, available at: http://dx.doi.org/10.1016/j.mbs.2006.08.012

[39] BLÜTHGEN N., BRUGGEMAN F.J., LEGEWIE S., HERZEL H., WESTERHOFF H.V., KHOLODENKO B.N.: 'Effects of sequestration on signal transduction cascades', *FEBS J.*, 2006, **273**, pp. 895–906

[40] FEINBERG M.: 'Chemical reaction network structure and the stability of complex isothermal reactors I. the deficiency zero and deficiency one theorems', *Chem. Eng. Sci.*, 1987, **42**, (10), pp. 2229–2268

[41] FEINBERG M.: 'Chemical reaction network structure and the stability of complex isothermal reactors II. Multiple steady states for networks of deficiency one', *Chem. Eng. Sci.*, 1988, **43**, (1), pp. 1–25

[42] BLÜTHGEN N., HERZEL H.: 'How robust are switches in intracellular signaling cascades?', *J. theor. Biol.*, 2003, **225**, (3), pp. 293–300

[43] KLIPP E., LIEBERMEISTER W., WIERLING C.: 'Inferring dynamic properties of biochemical reaction networks from structural knowledge', *Genome Inform.*, 2004, **15**, (1), pp. 125–137

[44] KIM P.M., TIDOR B.: 'Limitations of quantitative gene regulation models: a case study', *Genome Res.*, 2003, **13**, (11), pp. 2391–2395

[45] EISSING T., ALLGOEWER F., BULLINGER E.: 'Robustness properties of apoptosis models with respect to parameter variations and intrinsic noise', *IEE Proc. Sys. Biol.*, 2005, **152**, pp. 221–228

[46] MAURYA M.R., SUBRAMANIAM S.: 'A kinetic model for calcium dynamics in RAW 264.7 Cells: 1. mechanisms, parameters and sub-populational variability', *Biophys.*, 2007, **93**, pp. 709–728

[47] STELLING J., SAUER U., SZALLASI Z., DOYLE F. JR., DOYLE J.: 'Robustness of cellular functions', *Cell*, 2004, **118**, (6), pp. 675–685

[48] CONRADI C., FLOCKERZI D., RAISCH J.: 'Multistationarity in the activation of a MAPK: parametrizing the relevant region in parameter space', *Math. Biosci.*, 2008, **211**, (1), pp. 105–131

[49] SIGAL A., MILO R., COHEN A., GEVA-ZATORSKY N., KLEIN Y., LIRON Y., *ET AL.*: 'Variability and memory of protein levels in human cells', *Nature*, 2006, available at: http://dx.doi.org/10.1038/nature05316

[50] RASER J.M., O'SHEA E.K.: 'Control of stochasticity in eukaryotic gene expression', *Science*, 2004, **304**, (5678), pp. 1811–1814

[51] NEWMAN J.R.S., GHAEMMAGHAMI S., IHMELS J., BRESLOW D.K., NOBLE M., DERISI J.L., *ET AL.*: 'Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise', *Nature*, 2006, **441**, (7095), pp. 840–846

[52] ELOWITZ M.B., LEVINE A.J., SIGGIA E.D., SWAIN P.S.: 'Stochastic Gene expression in a single cell', *Science*, 2002, 297, pp. 1183–1186

[53] SENSSE A.: 'Convex and toric geometry to analyze complex dynamics in chemical reaction systems', PhD Thesis, University of Magdeburg, 2005

[54] PEDRAZA J.M., VAN OUDENAARDEN A.: 'Noise propagation in gene networks', *Science*, 2005, **307**, (5717), pp. 1965–1969

[55] KAERN M., ELSTON T.C., BLAKE W.J., COLLINS J.J.: 'Stochasticity in gene expression: from theories to phenotypes', *Nat. Rev. Genet.*, 2005, **6**, (6), pp. 451–464

[56] HUANG C.F., FERRELL J.E.J.: 'Ultrasensitivity in the mitogen-activated protein kinase cascade', *Proc. Natl. Acad. Sci. USA*, 1996, **93**, (19), pp. 10078–10083

[57] GUNAWARDENA J.: 'Multisite protein phosphorylation makes a good threshold but can be a poor switch' *Proc. Natl. Acad. Sci., USA*, 2005, **102**, (41), pp. 14617–14622

[58] HEINRICH R., NEEL B.G., RAPOPORT T.A.: 'Mathematical models of protein kinase signal transduction', *Mol. Cell.*, 2002, **9**, (5), pp. 957–970

[59] KLAMT S., SAEZ-RODRIGUEZ J., GILLES E.D.: 'Structural and functional analysis of cellular networks with CellNetAnalyzer', *BMC Syst. Biolo.*, 2006, **1**, (2)

[60] Available at: http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html

[61] Feinberg M., Ellison P.R.:'The chemical reaction network toolbox; available at: http//www.chbmeng.ohio-state.edu/ feinberg/crnt

[62] Available at: http://www.mpi-magdeburg.mpg.de/projects/promot