

# Integration of Boolean models exemplified on hepatocyte signal transduction

Rebekka Schlatter\*, Nicole Philipp\*, Gaby Wangorsch, Robert Pick, Oliver Sawodny, Christoph Borner, Jens Timmer, Michael Ederer and Thomas Dandekar

Submitted: 14th August 2011; Received (in revised form): 22nd September 2011

## Abstract

The number of mathematical models for biological pathways is rapidly growing. In particular, Boolean modelling proved to be suited to describe large cellular signalling networks. Systems biology is at the threshold to holistic understanding of comprehensive networks. In order to reach this goal, connection and integration of existing models of parts of cellular networks into more comprehensive network models is necessary. We discuss model combination approaches for Boolean models. Boolean modelling is qualitative rather than quantitative and does not require detailed kinetic information. We show that these models are useful precursors for large-scale quantitative models and that they are comparatively easy to combine. We propose modelling standards for Boolean models as a prerequisite for smooth model integration. Using these standards, we demonstrate the coupling of two logical models on two different examples concerning cellular interactions in the liver. In the first example, we show the integration of two Boolean models of two cell types in order to describe their interaction. In the second example, we demonstrate the combination of two models describing different parts of the network of a single cell type. Combination of partial models into comprehensive network models will take systems biology to the next level of understanding. The combination of logical models facilitated by modelling standards is a valuable example for the next step towards this goal.

**Keywords:** Boolean modelling; model integration; pathway; apoptosis

Corresponding author. Thomas Dandekar. Department of Bioinformatics, Biocenter, Am Hubland, University of Wuerzburg, 97074 Wuerzburg. Tel: +49(0)931-318-4551; Fax: +49(0)931-318-4552; E-mail: dandekar@biozentrum.uni-wuerzburg.de

\*These authors contributed equally to this work.

**Rebekka Schlatter** is a doctoral student in Institute for System Dynamics, University of Stuttgart integrating systems biology with various engineering and modelling approaches

**Nicole Philipp** spent thesis and first postdoctoral stay in bioinformatics in Department of Bioinformatics, Biocenter, Am Hubland, University of Wuerzburg, Würzburg, combining genome-based bioinformatics with systems biological modelling and is now post-doc at the MPI for aging, Cologne

**Gaby Wangorsch** is a doctoral student of bioinformatics in Department of Bioinformatics, Biocenter, Am Hubland, University of Wuerzburg, Würzburg, combining genome-based bioinformatics with systems biological modelling

**Robert Pick** worked experimentally on cellular regulation including apoptosis and proliferation with Borner in Freiburg, is now postdoctoral at Walter Brendel Centre of Experimental Medicine, Ludwig-Maximilians-University Munich

**Oliver Sawodny** is professor and chair at the institute for systems dynamics in University of Stuttgart, Pfaffenwaldring Stuttgart

**Christoph Borner** is a professor and chair at Institute of Molecular Medicine and Cell Research (ZBMZ), as well as BIOSS Centre for Biological Signalling Studies, Albert Ludwigs University Freiburg, Freiburg

**Jens Timmer** directs the Centre for Data Analysis and Modelling, works further at BIOSS Centre for Biological Signalling Studies and is professor for theoretical physics in Institute of Physics and Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Germany

**Michael Ederer** is a group leader of the institute for systems dynamics in Institute for System Dynamics, University of Stuttgart, Pfaffenwaldring, Stuttgart

**Thomas Dandekar** is the head of the department of bioinformatics, Biocenter, Am Hubland, University of Wuerzburg Würzburg, combining genome-based bioinformatics with systems biological modelling

## INTRODUCTION

Dynamical models have the potential to provide a realistic description of cellular events as they are capable of reproducing time dependent processes in a quantitative manner. We adhere to the final goal of a comprehensive dynamic model of signal transduction processes. However, we are not able to reach this vision short term. Due to the large size of the cellular network in combination with experimental bottlenecks in parameter measurement, it is still unfeasible to build a comprehensive dynamic cell model. Instead, a stepwise but nevertheless, aim-oriented procedure is needed. We will propose here details towards this goal, using the liver and especially the hepatocyte as an example.

We will particularly focus, thereby, on the potential of Boolean or logical models. They are well established, e.g. in the field of gene expression [1]. Furthermore, they already allow predictions when kinetic data are sparse and detailed dynamic approaches such as modelling on differential equations, kinetic rate equations [2] or power laws (S-Systems) [3], suffer from insufficient data for the approach. Recently, more and more attention is being paid to the Boolean modelling of signal transduction [4–7]. The relevant literature is actually quite rich, and a closer look reveals tens of relevant articles. Thus Medline lists 153 articles (as of 15 September 2011) using the keywords Boolean model(l)ing with a clearly increasing trend to use such models. Groups working in this area include among others, Alvarez-Buylla and colleagues [8, 9] on dynamical properties of gene transcription networks. For their Boolean models, they used hundreds of microarray experiments to infer the nature of the regulatory interactions among genes, for *Saccharomyces cerevisiae*, *Escherichia coli* and *Bacillus subtilis*, *Drosophila melanogaster* and flower development of *Arabidopsis thaliana*. Discrete logical models are furthermore developed by Sorger and co-workers, for instance to compare normal and transformed hepatocytes [10]. Their approach specifically deals also with the fuzzy logic of biochemical signals [11]. Garg *et al.* [12] provided algorithms based on reduced ordered binary decision diagrams (ROBDDs) for Boolean modelling of gene regulatory networks of T helper cells and Th1–Th2 cellular differentiation. Boolean models are now easily transformed into continuous models (see algorithm comparisons below) as for instance Wittmann *et al.* [13] demonstrate in a concrete example on T-cell receptor signalling.

Recently, Wang and Albert [14] introduced elementary signalling modes to predict the essentiality of signal transduction network components and show strong agreement with the results of their Boolean (logic) dynamic models and experimental observations. Chaves *et al.* [15] compare Boolean and piece-wise affine differential models for genetic networks regarding the carbon starvation response network in *E. coli*. The comparison yields new tools for analysis and reduction of biological networks, robustness of Boolean networks and asynchronous Boolean dynamics. Larger and more complex gene regulatory networks come in reach by Boolean models such as the gene regulatory network underlying mammalian cortical area development [16].

We, hence, believe Boolean models to be currently the most adequate approach to reproduce comprehensive signal transduction networks (e.g. [17–19]). In order to reach this goal, one needs to integrate Boolean models of subsystems into more comprehensive models.

Model integration suffers from several pitfalls. Combined models generally are over-proportionally complex because the behaviour of the model components is context-dependent and the interactions between the submodels leads to non-trivial behaviour. For smooth model integration, it is essential that the submodels adhere to consistent modelling standards. Based on these standards, we will demonstrate two different ways of integration. We start with a model integrating two different cell types and their interactions and then illustrate the fusion of two models operating within one cell type.

The overall goal of a unifying dynamical and quantitative model providing holistic understanding of cellular signalling is the declared aim of many systems biology research projects. We suggest that the comprehensive description of cellular signalling can be reached by Boolean modelling and model integration. Comprehensive, qualitative Boolean models may then pave the way for the development of comprehensive, quantitative and dynamic models.

## Comparison of different approaches and tools for Boolean modelling

A number of tools allow Boolean or semi-quantitative modelling, for instance, Copasi (signalling example in Ref. [32]), S-Systems (including power law transformation or approximation [3])

and Petri Nets [33]. Specific software for Boolean-type of modelling includes the Genetic Network Analyzer (GNA [20]). GNA employs piece-wise linear (PL) differential equation models that have been well studied in mathematical biology. While abstracting from the precise molecular mechanisms involved, the PL models capture essential aspects of gene regulation. Their simple mathematical form permits a qualitative analysis of the dynamics of the genetic regulatory systems to be carried out. Instead of numerical values for parameters and initial conditions, GNA asks the user to specify the qualitative constraints on these values in the form of algebraic inequalities. Unlike precise numerical values, these constraints can usually be inferred from the experimental literature and incorporated as algebraic inequalities. Analysis of the state transition graphs by means of VisualGNA allows one to investigate in detail the predicted qualitative equilibrium state, as well as qualitative behaviours leading to the equilibrium state. Furthermore, Boolnet is an R package for generation, reconstruction and analysis of Boolean networks [21]. This is a powerful package and efficiently integrates methods for synchronous, asynchronous and probabilistic Boolean networks. This includes reconstructing networks from time series, generating random networks, robustness analysis via perturbation, Markov chain simulations and identification and visualization of attractors applying various R routines. Besides this, Albert *et al.* [22] present BooleanNet. This is a software library that can perform Boolean modelling simulations based on simple text inputs. Quite complex networks can be tackled, e.g. modelling abscisic acid (ABA)-induced stomatal closure in plants, T-cell large granular lymphocyte leukaemia simulation or modelling the mammalian immune response to *Bordetella bronchiseptica* infection. Synchronous and asynchronous updates of simulations are possible, as well as exploiting the piece-wise linear formalism to build specialized Boolean rules that can represent each individual node. For comparison, GINsim (Gene Interaction Network simulation) [23] is a computer tool for the modelling and simulation of genetic regulatory networks featuring a simulator of qualitative models of genetic regulatory networks based on a discrete, logical formalism. Structural analysis on the network is possible, e.g. elementary circuits, inclusion of perturbations (e.g. mutations), state transitions (calculated using priority classes for interactions, individual nodes get discrete maximal

level and basal level, incoming interactions and their logical connections).

A recent tool of interest is the Network-Free Stochastic Simulator (NFsim), a general-purpose modelling platform that overcomes the combinatorial nature of molecular interactions [24]. Instead of representing molecular species as variables in equations, NFsim uses a biologically intuitive representation: objects with binding and modification sites acted upon by reaction rules. Reaction rates can thus, be defined as arbitrary functions of molecular states to provide powerful coarse-graining capabilities, e.g. to merge Boolean and kinetic representations of biological networks. The authors demonstrate this with models of immune system signalling, microbial signalling, cytoskeletal assembly and oscillating gene expression.

As powerful, handy and easy to use Boolean modelling tools, we review more closely Standardized Qualitative Dynamical systems (SQUAD) [25] and CellNetAnalyzer (CNA) [26]. These two approaches are complementary to each other.

CNA is a Matlab toolbox for Boolean modelling and model analysis. A key analysis method is the computation of the partial logical steady state for given initial conditions. CNA computes the values of nodes that approach a unique steady state. Node values that show multiple steady states and oscillations, as they may occur in networks with feedback loops, are excluded from the steady state analysis. It is often possible to break feedback loops by omitting interactions based on biological considerations such that steady state analysis of all nodes becomes possible, see for example ref. [27]. Further, the CNA toolbox includes among others, the pathway and feedback analysis, automated search for minimal intervention and cut sets and calculation of dependency matrices.

The Java-based tool SQUAD uses a heuristic algorithm based on concatenated exponential functions to build a differential equation system from a given Boolean model. The differential equation system is supposed to provide a qualitative approximation of the transient network behaviour. Overall, SQUAD is fast, easy to use and is helpful to get a full and fast system overview for systematic comparisons.

SQUAD, as well as CNA calculate the (partial) logical steady state of a network but do not use any so-called updating strategy to simulate the network response. As described by Fauré *et al.* [28], updating strategies can be classified in synchronous

and asynchronous updating. However, both types of updating strategies require assumptions on the rate and order of the state transitions and thus require biological knowledge that is often not available. Alternatively, one can simulate the network for many times and perform statistical analyses [29]. These problems are circumvented by calculating the logical steady state of the network because it does not depend on the updating strategy. For these reasons, we recommend to build logical models of large-scale networks with a steady state-oriented approach for the present.

Besides these two tools and their respective approaches, similar other tools and approaches to Boolean modelling exist, e.g. Odefy [30] as a closely related approach uses a modelling technique called *HillCube* to convert a Boolean model into a model of differential equations [31]. A selection of features pertaining to this related method is compared with the two approaches intensively used in this article (Table 1).

Note that our tool comparison gives only a first idea about the differences and specific advantages and limitations of the various tools.

## BOOLEAN MODELLING OF APOPTOSIS IN HEPATOCYTES

Two examples of Boolean models describing apoptosis signalling in hepatocytes form the basis for further model integration. One study used SQUAD [5]. It samples over the different network states in a dynamical way, starting from a Boolean network that modelled Fas ligand (FasL)-mediated apoptosis in liver cells. A complex and internally strongly linked network was assembled around this (74 nodes, 108 edges) and a number of alternative crosstalk possibilities and networks were considered. Four robust and stable system states were identified: two states comprise cell survival and two describe apoptosis by the intrinsic or extrinsic pathway, respectively. The model (<http://boolean.bioapps.biozentrum.uni-wuerzburg.de/index.php>) was validated by comparing it with experimental data from kinetics of cytochrome c release and caspase activation in wild-type and bid knockout cells grown on different substrates. Pathophysiological model modifications produce output behaviour that agrees well with experimental data such as input from cytomegalovirus proteins M36 and M45. Intercellular

**Table I:** Comparison of technical features in SQUAD, Odefy and CNA

Feature	SQUAD	Odefy	CNA
Input format	net, mml and xml:sbml	Boolean formulas, yED, CNA, GINSim, PBN	CNA format, SBML
Output format	canonical time-series plot, VirtualBlot, numerical data in a plain text file with tab-separated values	SQUAD, GNA, MATLAB script files, SB toolbox, SBML and R script files	CNA, Matlab, text, SBML
Compatibility	Only with Celldesigner 3.5.I	Matlab—and octave—compatible	Matlab and to some extent octave
Version compared	Version 2.0	Version 2010	Version 2010
Open source	Yes	Yes	No (but free download)
System state analysis	Yes, strength of the method, rapid sampling over all system states, system state directly transferred to nodes.	Yes, but more time consuming, excel Table of nodes and their activation	Yes, but more time consuming, excel Table of nodes and their activation
Simulation	Basic simulation and visualization functionalities for the continuous models. The software permits to make simulations on the continuous system, allowing for the modification of several parameters.	Basic simulation and visualization functionalities for both the Boolean, as well as the continuous models. Multi-level logical states are implemented (2 or n times activated/inhibited)	Calculation of dependency matrix, paths and cycles, logical steady state analysis (I/O behaviour), determination of minimal intervention sets
Transformation	SQUAD converts the network into a discrete dynamical system, and it uses a binary decision diagram algorithm to identify all the steady states of the system. Then, the software creates a continuous dynamical system and localizes its steady states which are located near the steady states of the discrete system.	Automated transformation of Boolean equations or graphs or imported models into systems of ordinary differential equations by multivariate polynomial interpolation and optional application of sigmoid Hill functions.	Logical hypergraphs can be transformed into interaction graphs. Odefy can be used as plug-in (on the fly transformation to a dynamic model in ODEfy format)
Download link	<a href="http://www.enfin.org/wiki/doku.php?id=enfin:squad:start">http://www.enfin.org/wiki/doku.php?id=enfin:squad:start</a>	<a href="http://cmb.helmholtz-muenchen.de/odefy">http://cmb.helmholtz-muenchen.de/odefy</a>	<a href="http://www.mpi-magdeburg.mpg.de/projects/cna/download.html">http://www.mpi-magdeburg.mpg.de/projects/cna/download.html</a>

regular interactions considered below involve in addition TNF and IL6.

Schlatter *et al.* [4] built a literature-based, large-scale Boolean model of the central intrinsic and extrinsic apoptosis pathways, as well as pathways connected with them using CNA. The logical apoptosis model comprises 86 nodes and 125 interactions. The model responds to several external stimuli such as Fas ligand, TNF $\alpha$ , ultraviolet radiation B (UV-B) irradiation, interleukin-1 $\beta$  and insulin and is freely available. Extensions of classical Boolean models, namely timescales and multi-value node logic, were used in this study and shown to be indispensable to reproduce the behaviour of the apoptotic network. The coherence of the model with measurement data was extensively experimentally validated. Thereby, an UV-B dose effect is shown for the first time in mouse hepatocytes. The logical model of apoptosis provides valuable information about the topology of the network including feedback loops and crosstalk effects. Analysis of the model revealed a tight regulation emerging from high connectivity and spanning crosstalks and a particular importance of feedback loops. TNFR1-induced apoptosis (in contrast to Fas) is a two-step process, which involves two sequential signalling complexes. Regarding this, an unexpected feedback from Smac release to receptor-interacting protein (RIP) could further increase complex II formation (TNF receptor-associated death domain (TRADD) and RIP1 associate with Fas-Associated protein with Death Domain (FADD) and caspase-8, thereby forming a cytoplasmic complex). The introduced Boolean model provides a comprehensive and coherent description of the apoptosis network behaviour. It gives new insights into the complex interplay of pro- and anti-apoptotic factors and can be easily expanded to other signalling pathways.

## MODELLING STANDARDS FACILITATE DIRECT COMBINATION OF BOOLEAN MODELS

A direct combination of two mathematical models requires a compatible modelling approach. The combination of dynamical, differential equation models is quite challenging as the parameters of differential equation models often do not correspond to elementary kinetic parameters but subsume different biochemical processes, new parameter

estimates for the combined model might be needed. The combinatorial explosion of complexity makes parameter estimation in a detailed differential equation model of the whole liver cell impossible in the light of current technical possibilities. Combination of Boolean models describing different aspects of the same system is less complex although not trivial.

Consistent modelling standards are indispensable for successful combination of Boolean models regardless of which tool is used. To encourage model convergence towards a comprehensive overall Boolean model, e.g. for the liver cell and/or its cell-cell interactions, we propose the following conventions on node values (i and ii), quantitative experimental data (iii), time (iv), input and output (v and vi) and unknown components (vii).

(i) Definition of node values: Classical Boolean algebra only provides two node values namely 'on' and 'off' or '1' and '0'. However, many biological interactions cannot be realized with only one level of activation. For example, the decision about CD95-mediated apoptosis is based on typical threshold behaviour [34] and a dose-dependent effect of UV irradiation was shown in hepatocytes [4]. We propose to discretize the 'on' state of affected nodes either modelling several different nodes representing different active states or using multi-value logic as implemented in CNA. In fact, both approaches are compatible as also in CNA, the logical equations for each multi-level node status can be written separately. However, we propose that the 'off' state of every node must stay equivalent to the absence of the according chemical species, e.g. the knockout of the according gene. So any node in the 'off' state is not able to participate actively in an interaction according to our standards as a matter of principle. We call the definition of a node value by a certain functional property, a functional definition of the node value. Using this standardized functional definition of the node value 'zero', the knockout of a gene is equivalent to setting its node value to zero.

Regardless of using single or multi-value logic, we propose such a functional definition also for the 'on' state(s).

(ii) Boolean models are not quantitative and we do not recommend the attempt to imitate quantitative model behaviour; e.g. there should be no excessive use of multi-level values. Instead, the special potency of Boolean models for qualitative predictions can be

strengthened by functional definition of node values. A discretization of the ‘on’ state may only be introduced and accordingly assigned if there are functionally different effects on the network related to the respective node, e.g. if a high, low and zero (or very low) concentrations of a protein have qualitatively different effects.

(iii) Quantitative experimental data: this issue is directly related to the topic of experimental validation of Boolean models. Measurement values usually consist of information on quantities (concentration, activity, etc.) plus an according point of time. Both aspects cannot be directly translated to a logical model. Regarding quantity, we propose to check for functional dependencies. If a node A is defined to be ‘1’ by its function to activate node B, the activation of B should also be measured in order to prove ‘A=1’ in a certain setting.

(iv) Treatment of time: as time is not respected while calculating the logical steady state of a Boolean network, this information seems to be dispensable. However, the measurement time point is of course not arbitrary. A major focus of Boolean modelling is the description of the short-term response of signal transduction networks upon a stimulus or a set of stimuli. In such examples, one is mainly interested in the question whether a certain species, e.g. protein, is active at any time point during the signalling process. In such cases, it is appropriate to use the peak concentration over a time course. Other applications of Boolean modelling may require the formulation of similar standards.

(v) Output: finally, we discuss the possibility of introducing artificial nodes in a Boolean model which do not correspond to a single chemical species. Artificial nodes are useful to sum up the specific network response of interest such as ‘apoptosis’ or ‘survival’ [4, 5]. All pre-conditions for apoptosis included in the model can thereby be linked using logical gates. This approach has not only the advantage of making the model outcome visible at first sight, but also allows for automated analysis of, e.g. impact of a certain input node on the output node ‘apoptosis’ and thereby for the relevant biological question.

(vi) Input: besides the functional definition of output nodes, we also propose this option for input nodes. An artificial node termed “housekeeping” node can be used as a pre-condition to initialize the ‘on’ status of certain nodes in the model that are constitutively active in wild-type cells [4]. The

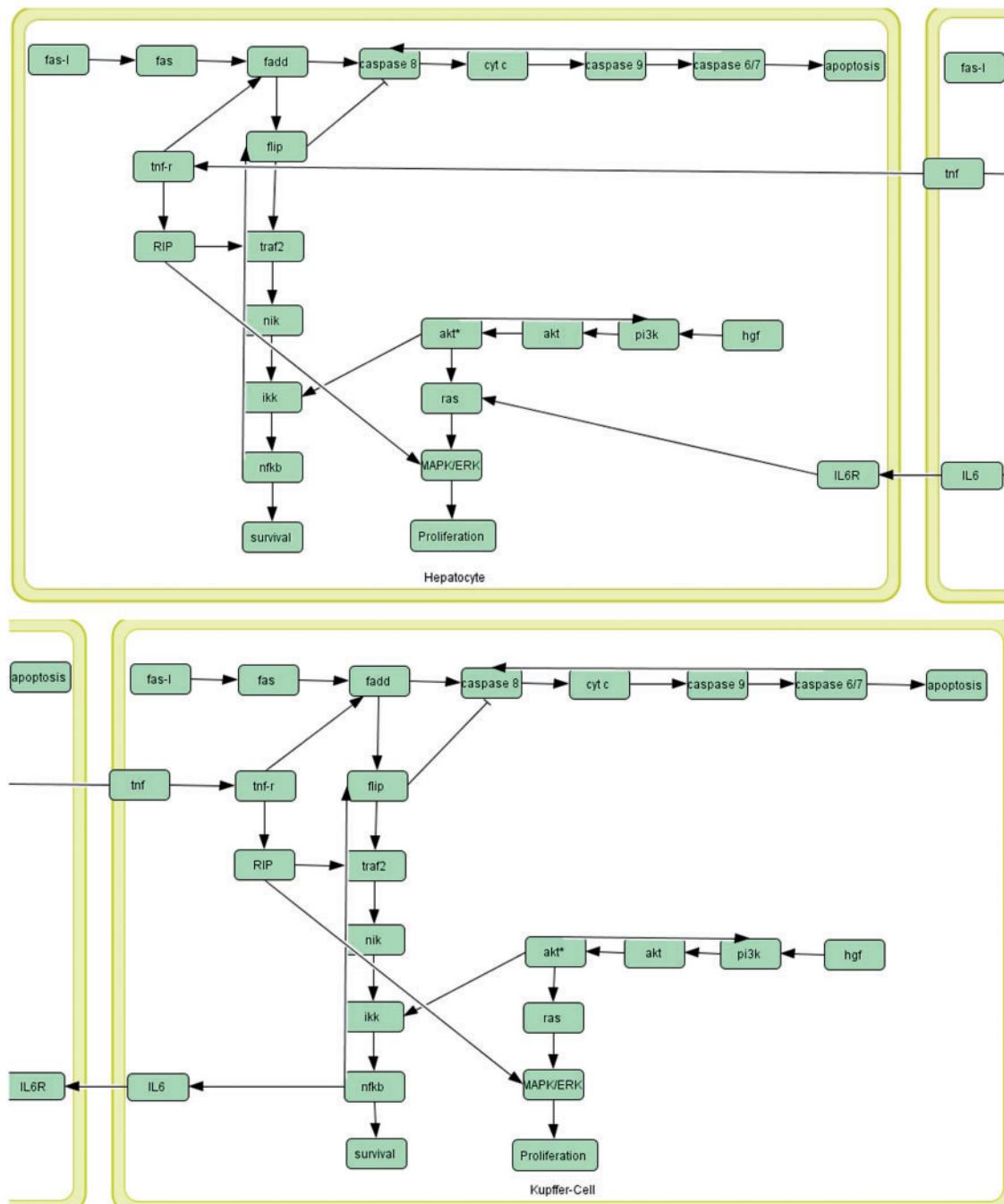
employment of such a ‘housekeeping’ node allows simulating the impact of transcriptional inhibitors such as Actinomycin D via setting the ‘housekeeping’ node to zero. Transcriptional inhibitors are here, drugs of special interest; they are not only used to investigate cellular networks experimentally but also to treat diseases such as cancer or HIV.

(vii) Treatment of unknowns: not least, artificial nodes can be used to model unknown interrelations and thereby to avoid doubtful assumptions on molecular interactions that could become independent otherwise during a model combination process. For example, it was shown that hepatocytes switch between type I and type II apoptosis depending on culturing conditions [35]. However, the underlying mechanism is not yet elucidated. The switch was modelled by [4] accordingly using an artificial node ‘P’ representing some unknown interaction(s) or protein(s). Artificial nodes like these are easy to recognize as such and can be replaced neatly as soon as the according cellular functionality has been clarified.

## COMBINATION OF BOOLEAN MODELS

### Combination of mathematical models establishing a cell–cell interaction

As a case study from our current work, we demonstrate the direct unification of a Boolean model of the hepatocyte with a Boolean model of the Kupffer cell by considering aspects of their cell–cell interaction (Figure 1). The SQUAD model introduced above forms the basis of the integrated model. First, both cell types are described by a similar model of the FasL, TNF $\alpha$  and IL-6 signalling pathways (hepatocyte version: Supplementary Figure S1). Here, the focus lies on the Fas-induced extrinsic apoptotic pathway, the NF $\kappa$ B correlated survival pathway and on the Ras-/MAPK-dependent proliferative pathway. The integrated model includes two connection pathways between hepatocytes and Kupffer cells describing their interplay. The activation of the NF $\kappa$ B pathway in Kupffer cells goes along with a secretion of IL-6. IL-6 itself binds and activates the IL-6 receptor in hepatocytes thus, activating the Ras-/MAPK-dependent proliferative pathway. The other interplay between hepatocytes and Kupffer cells deals with the secretion of TNF $\alpha$  from Kupffer cells, activating the TNFR-1-dependent TNF $\alpha$  mediated anti-apoptotic signalling in hepatocytes.



**Figure 1:** Network topology for the cell–cell interaction model. Apoptosis, proliferation and survival pathways belonging to the hepatocyte cell are shown on the top. Pathways for Kupffer cells are shown on the bottom with two interactions connecting both cell types. Method: The SBML file/model was set up with CellDesigner Version 3.5.1 and analysed with SQUAD on a windows computer.

The integrated cell–cell interaction model has four different stable steady states (Table 2).

In steady state 1 and steady state 2, all nodes of the survival pathway as well as proliferative nodes are active, whereas all apoptotic nodes are inactive. The difference between them is the activation of the proliferative pathway. In steady state 1, all nodes in

the model except the caspases are active and the hepatocyte gets its proliferative signal from stimulation with HGF. In steady state 2, Akt, as well as PI3K are inactive and the hepatocyte gets its proliferation signal via IL-6 from the Kupffer cell. In contrast, in steady state 3, only the apoptotic nodes are active whereas survival and proliferative nodes are inactive.

**Table 2:** Steady state analysis of the integrated cell–cell interaction model

Steady state	1	2	3	4	Pathway
Apoptosis	0	0	1	0	Apoptosis
Caspase6/7	0	0	1	0	Apoptosis
Caspase8	0	0	0.93	0	Apoptosis
Caspase9	0	0	1	0	Apoptosis
CytC	0	0	1	0	Apoptosis
IL6	1	1	0	0	Inflammation
IL6R	1	1	0	0	Inflammation
MAPK/ERK	0.93	0.91	0	0	Proliferation
Proliferation	1	1	0	0	Proliferation
AKT	1	0	0	0	Cytoskeleton
AKT*	1	0	0	0	Cytoskeleton
FLIP	0.93	0.93	0	0	Apoptosis
IKK	1	0.93	0	0	Inflammation
NFkB	1	1	0	0	Inflammation
NIK	1	1	0	0	Inflammation
PI3K	0.93	0	0	0	Phosphatidylinositol
RAS	1	0.93	0	0	Proliferation
Survival	1	1	0	0	Survival
TRAF2	0.91	0.91	0	0	Apoptosis

All key nodes of the cell–cell interaction model network are listed together with the activation state for each steady state. Full activation is represented by 1; inactive nodes are set to 0. Pathway classification is only for demonstration issues; in fact, it is difficult to differentiate exactly between the pathways.

In steady state 4, all nodes are inactive. This situation guarantees stable population of the liver as proliferation and apoptosis are each in cell-type equilibrations adjusted by the overarching hierarchical interactions.

The model is available from our website in XML format (<http://boolean.bioapps.biozentrum.uni-wuerzburg.de/index.php>). The complete cell-cell interaction model allows to be quickly assembled as the interactions are directly connecting the individual models and the comprehensive model integrates inter- and intracellular interactions: we are able to have a close look at two different cell types responding with apoptosis (example Supplementary Figure S2a) or proliferation and how IL-6 (Supplementary Figure S2b) or TNF (Supplementary Figure S2c) may influence this interaction in the cell–cell interaction model. Four different system states describe here the essence of the resulting system behaviour composed of the two cellular networks and their key interactions.

It is of course a strong simplification to focus on just two intercellular interactions and one must carefully evaluate the validity of the model for a given application. However, current experimental data available from this system are compatible with the model behaviour.

Thus, the model predicts well the behaviour of hepatocytes under different cultivation conditions

regarding viability and apoptosis, intrinsic and extrinsic pathway [5] and is compatible with further protein expression data regarding both pathways [30]. Moreover, the two key interactions chosen were found to be important in cell co-cultures of hepatocytes and Kupffer cells [36]. Thus, the demonstrated model integration strategy is a valuable tool to get insight into the mutual interaction of two cell types.

### Integration of mathematical models of different cellular pathways

As an example for the combination of two different signalling networks in one cell type, we present the integration of a comprehensive logical apoptosis model (Supplementary Figure S3) [4] with an apoptosis execution model presented here for the first time. The logical apoptosis model comprises 86 nodes and 125 interactions. It responds to several external stimuli such as Fas ligand, TNF $\alpha$ , UV-B irradiation, interleukin-1 $\beta$  and insulin. The coherence of the model (predicted apoptosis output) was experimentally validated [4]. The model focuses on the different signalling pathways leading to caspase-3 activation and their crosstalks. The activation of executioner caspase-3 was taken as indicator for apoptosis and in a simplified manner only four nodes (gelsolin, PARP, ICAD, CAD) have been

modelled between caspase-3 and the apoptosis output node.

In contrast, the apoptosis execution model focuses on the events following the activation of executioner caspases which lead to the final demise of the cell (Supplementary Figure S4). It comprises 60 nodes and 91 interactions. Input nodes of this model are caspase-3 and -7 and the output node is apoptosis. The model includes six artificial nodes describing central events in cell demise. These describe the nucleus DNA fragmentation, chromatin condensation and core fragmentation, as well as cellular condensation, blebbing and release. In the Supplementary Data, all nodes and logical equations of the apoptosis execution model are listed including the according literature references. Overall, the apoptosis execution model provides a solid composition of current knowledge on the final steps to cell death.

Both models have been designed with CNA thereby, using the logical steady state approach and have been modelled under consideration of the modelling standards described above. Therefore, the same understanding of artificial nodes and functional node value definition can be assumed. For example, both models employ a housekeeping node and the 'off' state of every node representing a protein is equivalent to the knockout of the according gene. There are also multi-value nodes in both models. For the logical apoptosis model, they are discussed in [4]. Amongst others, caspase-3 p17 is defined as apoptotic for node value '2'. The apoptosis execution model adopts this definition. Another example for multi-value logic in the apoptosis execution model is ICAD which also nicely illustrates the principle of functional node definition. The protein ICAD is necessary for correct folding of the protein CAD. Further, CAD is activated by cleavage of ICAD via active caspase-3. Therefore,  $ICAD=0$  which is equivalent to ICAD knockout leads to  $CAD=0$  because CAD does not fold properly. The 'on' state of ICAD needs to be discretized so that  $ICAD=1$  corresponds to the uncleaved ICAD and  $ICAD=2$  to the cleaved protein which is the only setting leading to CAD activation.

The complete model is provided as Supplementary Data. It comprises 136 nodes and 206 interactions, implying that 10 nodes and 5 interactions have been cut during the model combination process that will be discussed below. The complete network map of the unified model is shown in Supplementary Figure S5. A scheme of the unified model is shown

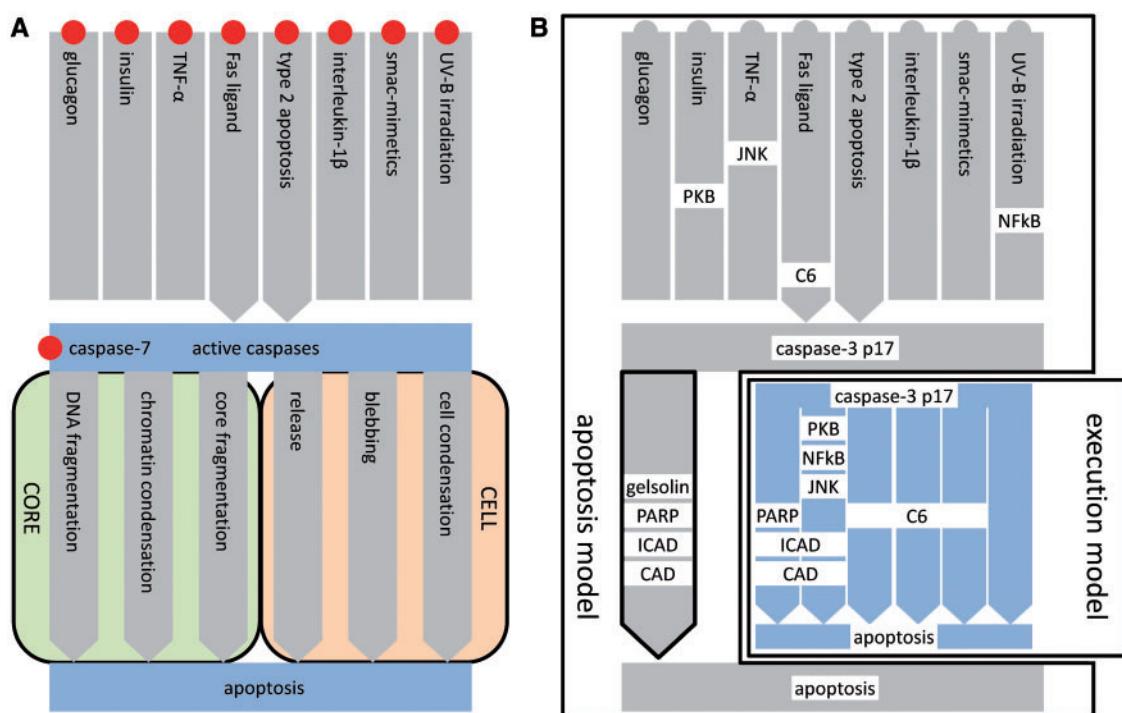
in Figure 2A which mainly represents the apoptosis model [4] and the Figure 2B corresponds to the apoptosis execution model. The two models are not only connected by handover of the caspase-3 node value, but both models share several common nodes making integration necessary.

We now describe the integration process in detail. In a first step, the defined species and interactions of both models have been pooled in common model files as basis for the combined model. In the second step, we have to decide about the handling of every single common node of the two initial models which are sketched in the model scheme in Figure 2B. The five nodes gelsolin, PARP, ICAD, CAD and apoptosis in the apoptosis model and their corresponding interactions as a whole have been replaced by the more detailed description in the execution model. In the combined model, the interactions for the nodes caspase-3 p17, caspase-6, JNK, NF $\kappa$ B and PKB are the union of the interactions of the single model with a single exception. The inhibition of the NF $\kappa$ B node via active caspase-3 p17 was excluded from logical steady state analysis because this interaction led to a feedback loop and no unique steady state exists. Experimental validation is needed to clarify whether this inhibition is relevant in hepatocytes. In addition, both models include a housekeeping node, which can be merged without difficulty as this is structurally an input node.

At this point of the procedure, the logical model has a unique steady state for a defined stimulus. However, because model integration may introduce unforeseen effects, we must ensure that the model still reproduces the same effects as the single models were required to do. The detailed analysis of occurring differences is necessary and can reveal new model predictions that need to be experimentally studied and thus, may give new insight in the system.

In order to analyse the combined model and to compare it with the single models, we applied quality assurance methods for Boolean modelling [27]. Exhaustive input testing allows for automated complete evaluation of all possible input settings. We checked all possible input node value combinations, 576 test settings in total. In every setting, we compared the values of the nodes for apoptosis, PKB, JNK, PARP, ICAD, CAD, caspase-6, caspase-3 p17 and NF $\kappa$ B from the apoptosis model [4] and the combined new model.

Both models result in the same results for the nodes PARP, ICAD, CAD, caspase-3 p17 and



**Figure 2:** Combined cellular pathway model. **(A)** Schema of the combined model with inputs (dark grey dots), signalling pathways (grey blocks) and the output apoptosis. **(B)** Schema of connecting nodes (white) between the apoptosis model (grey) and the execution model (blue; within black framed box at the bottom on the right) before model combination. The black framed path at the bottom left is replaced by the execution model (bottom right) during model combination.

NF $\kappa$ B. However, the node values were different for caspase-6 in 57.6%, for PKB in 29.2%, for JNK in 27.1% and for apoptosis in 19.4% of the test settings. The changed results for the caspase-6 node are due to the fact that its evaluation was excluded from logical steady state computation in the apoptosis model [4] but in the combined model activation by active caspase-3 p17 is included. The PKB node in the apoptosis model [4] is activated only by the insulin signalling pathway. In the new combined model, it is also influenced by active caspase-3 p17 in accordance with our biological knowledge. Therefore, we consider the changed behaviour of the model with respect to caspase-6 and PKB to be more reliable and take the new results as new predictions for future experimental validation.

The discrepancy in 19.4% of the test cases for the apoptosis node could be assigned to the influence of the JNK node. The JNK node in the apoptosis model [4] is a signal sink only activated by TNF and inhibited by NF $\kappa$ B. In the combined model, its position is much more central and its status is influenced by several additional players

(Supplementary Figure S5). In particular, JNK promotes chromatin condensation and is therefore pro-apoptotic. This circumstance allowed induction of apoptosis by a TNF stimulus alone which is wrong regarding hepatocytes. To correct the model output, we changed the interaction [(TNF activates JNK) or (no NF $\kappa$ B activates JNK)] to [(TNF activates JNK) and (no NF $\kappa$ B activates JNK)]. After this adjustment the exhaustive input testing was repeated and the results for the apoptosis node value are now the same as for the primary models in all test cases. In addition, discrepancy for the JNK node value increased in 46.5% of the settings which is considered as a shift to more correct results.

Overall, the direct combination of both models was achieved without fundamental adjustments and the complexity was only moderately increased. However, we took high advantage from the fact that both models were based on the same standards. Even a single interaction can make a huge difference in a strongly cross-linked network and we consider a complete model check as indispensable for integration and therefore recommend the application of quality assurance methods. An implementation

of the tools suggested earlier [27] can be freely downloaded as CNA plug-in.

## DISCUSSION

Boolean models allow a comprehensive understanding of the complex signalling network governing the fate of liver cells. They do not depend on detailed kinetic information but are established from the network topology. Furthermore, the analysis of the systems response to different combined inputs is a special potency of this approach. The impact of crosstalks is of outstanding importance for the resulting behaviour of the cell. Moreover, knockout and knock-in effects can easily be modelled as well as different compartments. Here, we demonstrate that Boolean models of subsystems can be connected to more comprehensive models when following common modelling standards. We believe that this approach has the potential to be expanded towards a comprehensive cell-covering model structure. Critically evaluated and validated, comprehensive Boolean models can serve as textures for dynamic, differential equation-based modelling by mapping crosstalks and contributing structural information. Thereby, Boolean models support the overall long-term goal of a dynamic whole cell model.

The here mentioned Boolean models gave new insights into liver signalling and promote our knowledge about hepatocytes. There are also promising clinical application aspects. For example, viral apoptosis blocking was demonstrated for M36 and M45 cytomegaly virus proteins [5] and a UV-B light dose effect in hepatocytes was shown [4]. Furthermore, infection by cytomegaly virus, as well as oncolysis by vaccinia virus have been tested [37] and are currently explored further. Comprehensive Boolean models can contribute to important biological research questions. Boolean models may help to reveal the changes induced by cultivation of cells under different conditions, e.g. collagen or suspension, as this modifies cellular properties and differentiation. Cell-cell interaction models can furthermore be applied for tissue engineering [38] including modified AKT signalling in the different cell types. Larger logical models also support development of new drugs. For example, SMAC-mimetics are used as anti-cancer treatment and are included in the used apoptosis model of Schlatter *et al.* [27] and the integrated model presented here. In addition, minimal intervention sets as implemented in CNA allow

target search in such models. Viral infections of liver cells can be better understood starting from Boolean cellular models [5] but now including specific cell-cell interactions. This includes therapeutic approaches such as the action of an oncolytic virus [37] on liver tissue plus metastases. Combined Boolean models generate semi-quantitative data on all involved nodes immediately, if desired even cell-type specific and including key information on qualitative changes in the system.

## SUPPLEMENTARY DATA

Supplementary data are available online at <http://bib.oxfordjournals.org/>.

### Key Points

- Integration of existing models is critical for a holistic or comprehensive understanding of cells.
- Boolean modelling does not require detailed kinetic information and are useful precursors for large-scale quantitative models. Boolean models are comparatively easy to combine.
- We propose modelling standards for Boolean models as a pre-requisite for smooth model integration.
- We demonstrate coupling of logical models for two cell types, as well as combination of two models describing different parts of the network of a single cell type.

### Acknowledgements

We thank all our colleagues who participated in lively discussions about the pros, cons and integration of Boolean models.

## FUNDING

The German Federal Ministry of Education and Research (BMBF): Virtual Liver Network (<http://www.virtual-liver.de>) grant number 0315751; the German Research Foundation DFG (Da 208/12-1); the MedSys SARA grant (grant numbers 0315395B, 0315395E); the Land Bavaria.

## References

1. Damiani C, Serra R, Villani M, *et al.* Cell-cell interaction and diversity of emergent behaviours. *IET Syst Biol* 2011;5: 137.
2. Millat T, Wolkenhauer O, Fischer RJ, *et al.* Modeling of cellular processes: methods, data, and requirements. *Methods Mol Biol* 2011;696:429–37.
3. Voit EO. Modelling metabolic networks using power-laws and S-systems. *Essays Biochem* 2008;45:29–40.

4. Schlatter R, Schmich K, Avalos Vizcarra I, et al. ON/OFF and Beyond – a Boolean model of apoptosis. *PLoS Comput Biol* 2009;5:e1000595.
5. Philippi N, Walter D, Schlatter R, et al. Modeling system states in liver cells: survival, apoptosis and their modifications in response to viral infection. *BMC Syst Biol* 2009;3:97.
6. Samaga R, Saez-Rodriguez J, Alexopoulos LG, et al. The logic of EGFR/ErbB signalling: theoretical properties and analysis of high-throughput data. *PLoS Comput Biol* 2009;5:e1000438.
7. Thakar J, Albert R. Boolean models of within-host immune interactions. *Curr Opin Microbiol* 2010;13(3):377–81.
8. Balleza E, Alvarez-Buylla ER, Chaos A, et al. Critical dynamics in genetic regulatory networks: examples from four kingdoms. *PLoS One* 2008;3(6):e2456.
9. Alvarez-Buylla ER, Chaos A, Aldana M, et al. Floral morphogenesis: stochastic explorations of a gene network epigenetic landscape. *PLoS One* 2008;3(11):e3626.
10. Saez-Rodriguez J, Alexopoulos LG, Zhang M, et al. Comparing signaling networks between normal and transformed hepatocytes using discrete logical models. *Cancer Res* 2011;71(16):5400–11.
11. Morris MK, Saez-Rodriguez J, Clarke DC, et al. Training signaling pathway maps to biochemical data with constrained fuzzy logic: quantitative analysis of liver cell responses to inflammatory stimuli. *PLoS Comput Biol* 2011;7(3):e1001099.
12. Garg A, Di Cara A, Xenarios I, et al. Synchronous versus asynchronous modelling of gene regulatory networks. *Bioinformatics* 2008;24(17):1917–25.
13. Wittmann DM, Krumsiek J, Saez-Rodriguez J, et al. Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. *BMC Syst Biol* 2009;3:98.
14. Wang RS, Albert R. Elementary signaling modes predict the essentiality of signal transduction network components. *BMC Syst Biol* 2011;5:44.
15. Chaves M, Tournier L, Gouzé JL. Comparing Boolean and piecewise affine differential models for genetic networks. *Acta Biotheor* 2010;58(2–3):217–32.
16. Giacomantonio CE, Goodhill GJ. A Boolean model of the gene regulatory network underlying Mammalian cortical area development. *PLoS Comput Biol* 2010;6(9):e1000936.
17. Veliz-Cuba A, Stigler B. Boolean models can explain bistability in the lac operon. *J Comput Biol* 2011;18:783–94.
18. Wang H, Lu HH, Chueh TH. Constructing biological pathways by a two-step counting approach. *PLoS One* 2011;6:e20074.
19. Jack J, Wambaugh JF, Shah I. Simulating quantitative cellular responses using asynchronous threshold Boolean network ensembles. *BMC Syst Biol* 2011;5:109.
20. De Jong H, Geiselmann J, Hernandez C, et al. Genetic Network Analyzer: qualitative simulation of genetic regulatory networks. *Bioinformatics* 2003;19(3):336–44.
21. Müssel C, Hopfensitz M, Kestler HA. BoolNet—an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics* 2010;26(10):1378–80.
22. Albert I, Thakar J, Li S, et al. Boolean network simulations for life scientists. *Source Code Biol Med* 2008;3:16.
23. Naldi A, Berenguier D, Fauré A, et al. Logical modelling of regulatory networks with GINsim 2.3. *Biosystems* 2009;97(2):134–9.
24. Sneddon MW, Faeder JR, Emonet T. Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nat Methods* 2011;8(2):177–83.
25. Di Cara A, Garg A, De Micheli G, et al. Dynamic simulation of regulatory networks using SQUAD. *BMC Bioinformatics* 2007;8:462.
26. Klamt SJ, Saez-Rodriguez JA, Lindquist L, et al. A methodology for the structural and functional analysis of signalling and regulatory networks. *BMC Bioinformatics* 2006;7:56.
27. Schlatter R, Knies D, Ederer M, et al. Analysis of Boolean Models using quality assurance methods from software engineering. *IEEE Multi-Conference on Systems and Control (MSC), Yokohama* 2010.
28. Fauré A, Naldi A, Chaouiya C, et al. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *BMC Bioinformatics* 2006;22:e124–31.
29. Maiwald T, Timmer J. Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics* 2008;24:2037–43.
30. Krumsiek J, Pölsterl S, Wittmann DM, et al. Odefy—from discrete to continuous models. *BMC Bioinformatics* 2010;11:233.
31. Wittmann DM, Krumsiek J, Saez-Rodriguez J, et al. Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. *BMC Syst Biol* 2009;3:98.
32. Kühnel M, Mayorga LS, Dandekar T, et al. Modelling phagosomal lipid networks that regulate actin assembly. *BMC Syst Biol* 2008;2:107.
33. Li C, Nagasaki M, Koh CH, et al. Online model checking approach based parameter estimation to a neuronal fate decision simulation model in *Caenorhabditis elegans* with hybrid functional Petri net with extension. *Mol Biosyst* 2011;7:1576–92.
34. Bentele M, Lavrik I, Ulrich M, et al. Mathematical modelling reveals threshold mechanism in CD95-induced apoptosis. *J Cell Biol* 2004;166:839–51.
35. Walter D, Schmich K, Vogel S, et al. Switch from type II to I Fas/CD95 death signaling on in vitro culturing of primary hepatocytes. *Hepatology* 2009;48:1942–53.
36. Schmich K, Schlatter R, Corazza N, et al. Tumor necrosis factor  $\alpha$  sensitizes primary murine hepatocytes to Fas/CD95-induced apoptosis in a Bim- and Bid-dependent manner. *Hepatology* 2011;53:282–92.
37. Zhang Q, Liang C, Yu YA, et al. The highly attenuated oncolytic recombinant vaccinia virus GLV-1h68: comparative genomic features and the contribution of F14.5L inactivation. *Mol Genet Genomics* 2009;282:417–35.
38. Schanz J, Pusch J, Hansmann J, et al. Vascularised human tissue models: a new approach for the refinement of biomedical research. *J Biotechnol* 2010;148:56–63.