Integrating time-series data in large-scale discrete cell-based models

Louis Fippo Fitime[1], Christian Schuster[2], Peter Angel[2], Olivier Roux[1], and Carito Guziolowski[1]

[1] LUNAM Université, École Centrale de Nantes, IRCCyN UMR CNRS 6597
(Institut de Recherche en Communications et Cybernétique de Nantes)
1 rue de la Noë – B.P. 92101 – 44321 Nantes Cedex 3, France.
louis.fippo-fitime@irccyn.ec-nantes.fr
http://www.irccyn.ec-nantes.fr/en/
[2] Division of Signal Transduction and Growth Control (A100), DKFZ-ZMBH Alliance,
Deutsches Krebsforschungszentrum, Heidelberg, Germany

Abstract. In this work we propose an automatic way of generating and verifying formal hybrid models of signaling and transcriptional events, gathered in large-scale regulatory networks.This is done by integrating temporal and stochastic aspects of the expression of some biological components. The hybrid approach lies in the fact that measurements take into account both times of lengthening phases and discrete switches between them. The model proposed is based on a real case study of keratinocytes differentiation, in which gene time-series data was generated upon Calcium stimulation.
To achieve this we rely on the Process Hitting (PH) formalism that was designed to consider large-scale system analysis. We first propose an automatic way of detecting and translating biological motifs from the PID database to the PH formalism. Then, we propose a way of estimating temporal and stochastic parameters from time-series expression data of action on the PH. Simulations emphasize the interest of synchronizing concurrent events.

Keywords: time-series data, large-scale network, hybrid models, compositional approach, stochastic simulation

1 Introduction

The comprehension of the mechanisms involved in the regulation of a cell-based biological system is a fundamental issue. These mechanisms can be modeled as biological regulatory networks, which analysis requires to preliminary build a mathematical or computational model. By just considering qualitative regulatory effects between components, biologic regulatory networks depict fairly well biological systems, and can be built upon public repositories such as the Pathways Interaction Database [13] and hiPathDB [16] for human regulatory knowledge.

In this work we built an hybrid model of signaling and transcriptional events, gathered in large-scale regulatory networks, which stochastic simulation parameters were inferred from gene expression time-series data. The integration of time-series data in

dynamical models have been addressed separately by approaches that either: (a) focus first on modeling at small-scale the system and then on refining or improving it through the fitting with some data points, such as methods based on differential equations [15,1,9], (b) integrate in an efficient and complete fashion large-scale models and high-throughput data regardless from the system dynamics [3,8], or (c) fit dynamical data to middle-scale networks using an stochastic sampling of the space of behaviors and therefore without guarantee on finding global optima [6]. With this work we aim to fill the gaps between the aforecited methodologies.

In the context of modeling and analyzing stochastic and concurent biological systems various formalisms have been introduced such as Stochastic Petri Nets which is suitable for the representation of parallel systems [10]. They have been successfully applied in many areas; in particular, the specification of Petri Nets allows an accurate modeling of a wide range of systems including biological systems [4]. The major problem of Stochastic Petri Nets is that, generally, they do not lead to compact models. In addition, they do not provide results to deal with the state space explosion and are thus computationally expensive when modeling large-scale biological networks.

The Stochastic pi-calculus formalism was introduced by [12] and used in [7] for the modeling of biological systems. Stochastic pi-calculus has a rich expressivity and is well adapted for the use of compositional approach In this work we rely on this formalism through the Process Hitting (PH) framework [11], since it is especially useful for studying systems composed of biochemical interactions, and provides stochastic simulation as well as efficient static methods to analyse dynamical properties of the system. The PH framework uses qualitative and discrete information of the system without requiring enormous parameter estimation tasks for its stochastic simulation. This framework has been previously used to verify dynamical properties on biological systems without integrating high-throughput experimental data.

In this work we provide a method to build a time-series data integrated PH model and we evaluate the prediction power of this model concerning the simulatenously predicted traces of 12 gene components of the system upon system stimulation. The main results of this work are: (1) automatic generation of PH models integrating gene transcription and signaling events, with and without synchronisation of concurrent events, from the Pathways Interaction Database, (2) parameter estimation from time-series data and parameter integration in the PH model, and (3) comparison of the PH model predictions and experimental results. To illustrate our approach, we used a time-series dataset of keratinocytes cells, which shows the fluctuations of gene expression across time upon Calcium stimulation. This dataset was built to study keratinocytes differentiation, a time-dependent process in which the sequence of activation of signaling proteins is not yet completely understood. The method proposed in this paper remains general and can be applied to other case-studies.

2 Data and Methods

The general workflow for integrating time-series data in a PH model is depicted in Fig. 1, in the following sections we detail some of the workflow steps.

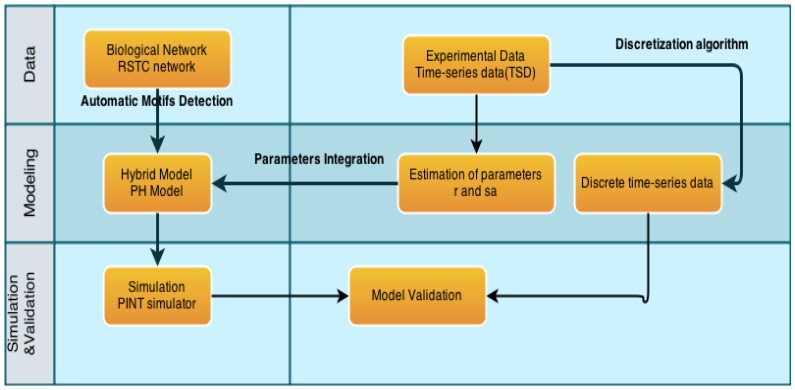


Fig. 1. Integrating stochastic and temporal information in a large-scale discrete biological model.

2.1 Data

Interaction graph. The interactions of the biological system under study were represented in an RSTC network, which stands for multi-layer receptor-signaling-transcription-cell state network and that was generated from the Pathway Interaction Database (PID). In order to build this network one needs to select a set of seed nodes related to the biological process studied. For our case study, the seed nodes were: (1) *E-cadherin*, which is a protein having Calcium binding domains and which plays an important role in cell adhesion; (2) the 12 significantly differentially expressed genes accross the 10 time-points; and (3) the cell states of keratinocytes-differentiation and cell-cycle-arrest. The network was extracted automatically from the whole content of the NCI-PID database by using a subgraph algorithm to link the seed nodes [2]. In Fig. 2 we show the RSTC network obtained.

Definition 1 (RSTC Network). *A* RSTC Network N *is a couple* (V, E)*, where:*

- $V = V_T \bigcup V_I$ *is the finite set of* nodes*; with* $V_T = \{v_{1t}, v_{2t}, \ldots, v_{n1t}\}$ *the set of terminal nodes;* $V_I = \{v_{1i}, v_{2i}, \ldots, v_{n2i}\}$ *the set of transient nodes.*
- $E = \{e_1, e_2, \ldots, e_m\}$ *is the set of edges.* $E \subseteq (V_T \times V_T) \bigcup (V_T \times V_I) \bigcup (V_I \times V_T)$

In this definition, terminal nodes can be genes, proteins, complexes, cellular states, biological processes and positive conditions. On the other side, transient nodes can be transcriptions, translocations, modifications and compounds. Edges are of different types: activation (agent), inhibition, output, input and protein-family-member.

Time-series microarray dataset. We use the time-series microarray data from Calcium stimulated human keratinocyte cells measured at 10 time-points. 200 transcripts were selected for their dynamic patterns, that is, their fold expression with respect to the non-stimulated cell was significant in at least one time point. We included in our model a subset of 12 of the 200 selected (see Fig. 3). These 12 genes had upstream regulatory

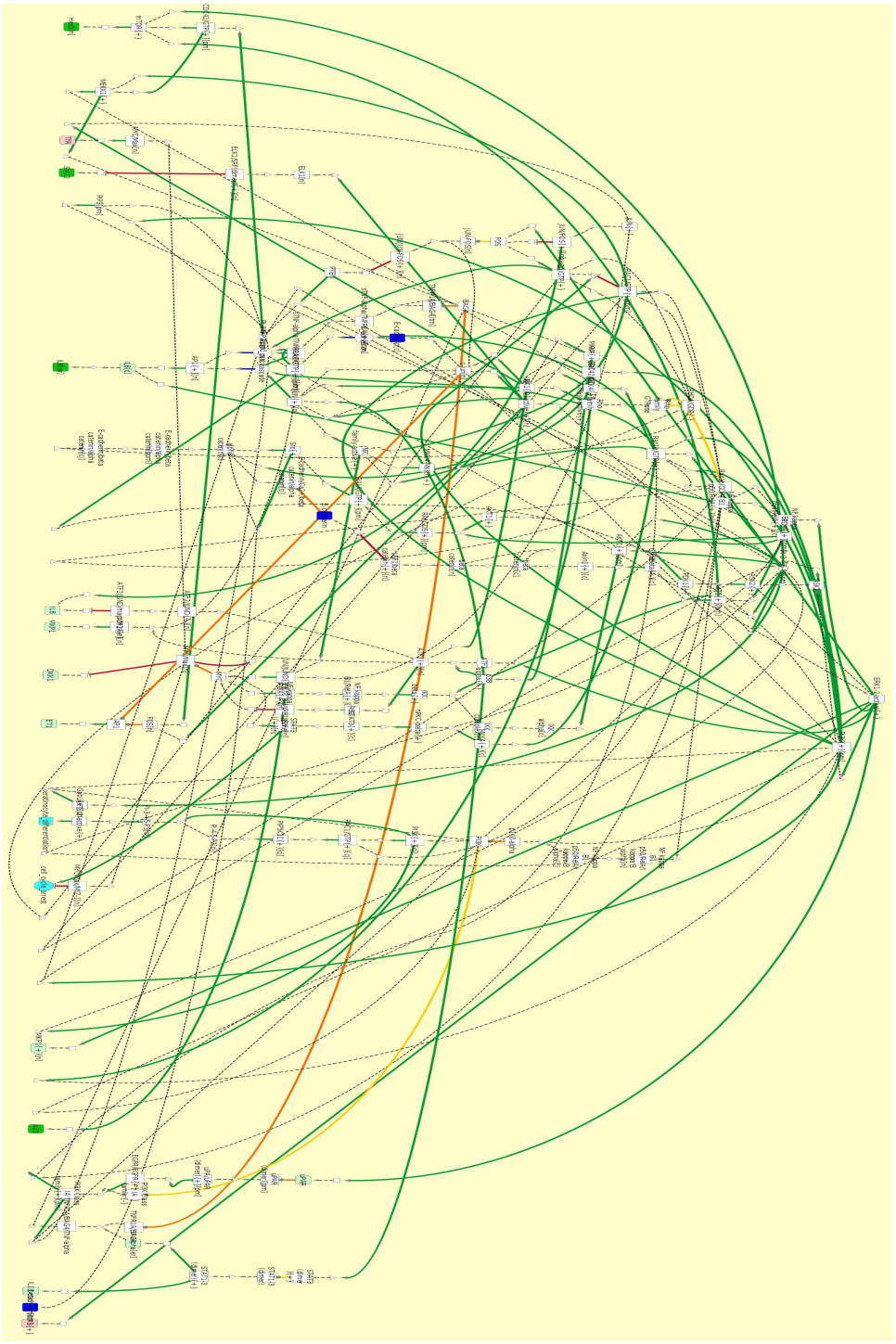


Fig. 2. Interaction graph linking E-cadherin with 12 genes of the time-series dataset. Blue nodes correspond to E-cadherin entities, red or green, to time-series genes, and cyan nodes to cellular processes. The graph is composed of 293 nodes and 375 edges (interactions). The set of nodes are composed of terminal nodes (proteins, complexes, genes, cellular state, biological processes and positive conditions) and of transient nodes (transcriptions, translocations, modifications and compounds). The set of edges are composed of interactions of type activation, inhibition, output, input and protein-family-member.

mechanisms when querying the PID-NCI database and therefore could be connected in the interaction graph.

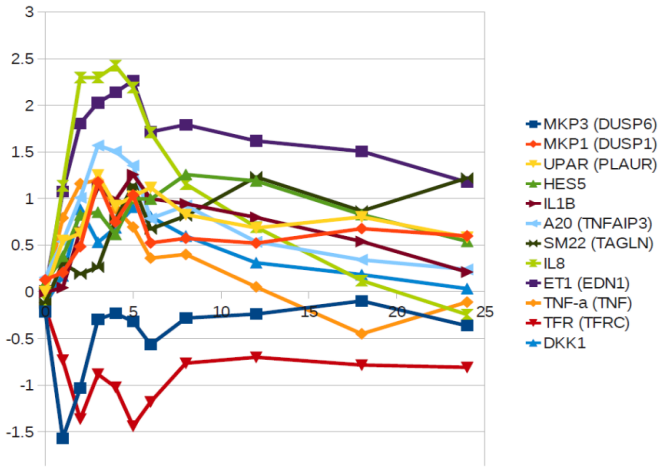


Fig. 3. 12 gene expressions upon Calcium stimulation.
The X axis represents time duration of the experiment measured in hours. The Y axis represents the log_2 expression level of genes with respect to control.

2.2 The Process Hitting Framework

Process Hitting (PH) gathers a finite number of concurrent processes grouped into a finite set of sorts. A sort stands for a component of a biological system while a process, which belongs to a unique sort, stands for one of its expression levels. At any time exactly one process of each sort is present. A state of the PH corresponds to such a set of processes. We denote here a process by a_i where a is the sort and i is the process identifier within the sort a. The concurrent interactions between processes are defined by a set of *actions*. Actions describe the replacement of a process by another of the same sort conditioned by the presence of at most one other process in the current state. An action is denoted by $a_i \to b_j \restriction b_k$, which is read as "$a_i$ *hits* b_j to make it bounce to b_k", where a_i, b_j, b_k are processes of sorts a and b, called respectively *hitter*, *target* and *bounce* of the action.

Definition 2 (Process Hitting). *A* Process Hitting *is a triple* $(\Sigma, L, \mathcal{H})$*, where:*

- $\Sigma = \{a, b, \dots\}$ *is the finite set of* sorts*;*
- $L = \prod_{a \in \Sigma} L_a$ *is the set of states with* $L_a = \{a_0, \dots, a_{l_a}\}$ *the finite set of* processes *of sort* $a \in \Sigma$ *and* l_a *a positive integer, with* $a \neq b \Rightarrow L_a \cap L_b = \emptyset$*;*
- $\mathcal{H} = \{a_i \to b_j \restriction b_k \in L_a \times L_b \times L_b \mid (a, b) \in \Sigma^2 \wedge b_j \neq b_k \wedge a = b \Rightarrow a_i = b_j\}$ *is the finite set of* actions*.*

Given a state $s \in L$, the process of sort $a \in \Sigma$ present in s is denoted by $s[a]$. An action $h = a_i \rightarrow b_j \nearrow b_k \in \mathcal{H}$ is *playable* in $s \in L$ if and only if $s[a] = a_i$ and $s[b] = b_j$. In such a case, $(s \cdot h)$ stands for the state resulting from playing the action h in s, with $(s \cdot h)[b] = b_k$ and $\forall c \in \Sigma, c \neq b, (s \cdot h)[c] = s[c]$. In order to model the fact that a molecule in the interaction graph is influenced by various molecules, two types of modeling-scenarios can be proposed: cooperation and synchronization.

Modeling cooperation The cooperation between processes to make another process bounce can be expressed in PH by building a *cooperative sort* [11]. Fig. 4 shows an example of a cooperative sort ab between sorts a and b, which is composed of 4 processes (one for each sub-state of the presence of processes in a and b). For the sake of clarity, processes of ab are indexed using the sub-state they represent. Hence, ab_{01} represents the sub-state $\langle a_0, b_1 \rangle$, and so on. Each process of sort a and b hits ab, which makes it bounce to the process reflecting the status of the sorts a and b (e.g., $a_1 \rightarrow ab_{00} \nearrow ab_{10}$ and $a_1 \rightarrow ab_{01} \nearrow ab_{11}$). Then, to represent the cooperation between processes a_1 and b_1, the process ab_{11} hits c_1 to make it bounce to c_2 instead of independent hits from a_1 and b_1. The same cooperative sort is used to make a_0 and b_0 cooperate to hit c_1 and make it bounce to c_0. Cooperation sort allows to model the fact that two components cooperate to hit another component.

Modeling synchronization The synchronization sort implements another type of cooperation. If we refer to the example of Fig. 4 left, we can similarly construct a *synchronization sort* ab between sorts a and b, defined with also 4 processes. Then, component c is activated (c_1 bounces to c_2 or c_0 bounces to c_1) if either a or b are activated. Therefore, each one of these processes ab_{01}, ab_{10}, ab_{11} can activate c. In order to inhibit c, both sorts, a and b, need to be in the sub-state 0, i.e. ab_{00}. Notice that this rule is a combination of OR logical gates for activation and AND logical gates for inhibition. Imposing the synchronization sort to model a target component regulated independently by multiple predecessors avoids oscillations in the behavior of the target component over time. These oscillations appear because each predecessor can independantly activate the target component when it is active, but when one predecessor is inhibited, it will inhibit the target component. This competition between the predecessors generates oscillations on the target component.

Example 1. Fig. 4 represents a PH $(\Sigma, L, \mathcal{H})$ with $\Sigma = \{a, b, c, ab\}$, and:

$$\begin{aligned} L_a &= \{a_0, a_1\}, \\ L_b &= \{b_0, b_1\}, \\ L_{ab} &= \{ab_{00}, ab_{01}, ab_{10}, ab_{11}\}, \\ L_c &= \{c_0, c_1, c_2\}. \end{aligned}$$

This example models a Biological Regulatory Network (BRN) where the component c has three qualitative levels, components a and b are Boolean and ab is a cooperative sort. In this BRN, ab inhibits c at level 2 through the cooperative sort ab (e.g. $ab_{00} \rightarrow c_2 \nearrow c_1$, $ab_{00} \rightarrow c_1 \nearrow c_0$) while a and b activate c through the cooperative sort ab (e.g.

$ab_{11} \rightarrow c_0 \not\uparrow c_1$ $ab_{11} \rightarrow c_1 \not\uparrow c_2$). Indeed, the reachability of c_2 and c_0 is conditioned by a cooperation of a and b as explained above.

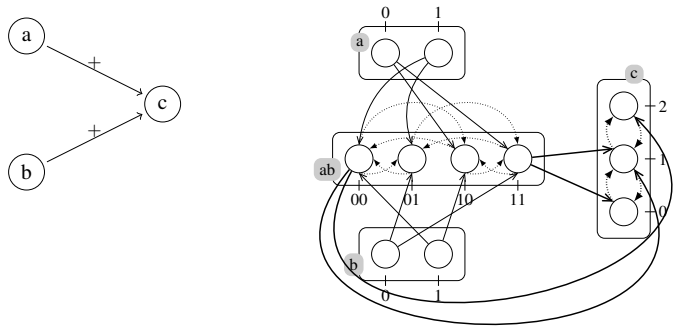


Fig. 4. (Left) biological pattern example. Nodes represent molecules (components) and edges, interactions. In this pattern components a and b cooperate to activate c. **(Right) equivalent PH model** with four sorts: three components (a, b and c) and a cooperative sort (ab). Actions targeting processes of c are drawn as thick lines.

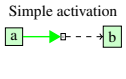
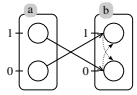
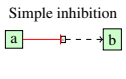
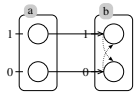
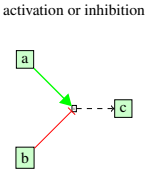
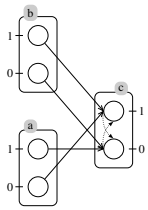
2.3 Model construction (from RSTC to PH)

Modeling the RSTC network as a PH model. In order to model the RSTC network as a PH model we select known biological regulatory patterns (atomic set of biological components and their interacting roles), represented as biochemical reactions in the RSTC network and propose their PH representation. Table 1 shows some examples of this transformation.

The automatic pattern selection and PH model generation algorithms use two procedures. The first takes as parameter argument the graph and automatically browses it node by node and detects all patterns in the graph. For each node (output node of the pattern) we call a recursive procedure, that allows to detect a minimal set of nodes (input node of the pattern) that has a direct influence over that node. This set of nodes plus the output node and the way input and output are linked form a pattern. The type of a pattern is determined by the type of the output node, the type of regulations that arrive to that node and the type of input nodes of the pattern. Consequently, the algorithm of patterns detection returns the pattern and its type to another procedure which translates the pattern into the PH formalism. This transformation takes care of different cases (cooperation, synchronization, simple activation, simple inhibition, etc.)

For example a molecule a cooperatting with a molecule b to activate a molecule c (Fig. 4, left), is a regulatory pattern because it is a protein-complex biochemical reaction that appears recurrent times. We model this pattern by four sorts (Fig. 4, right) a, b, c and ab. Sorts a, b and c stand for components a, b and c. The cooperative sort ab is introduced in order to characterize constraints on the components a and b. In the RSTC network, we find 25 regulatory patterns. We show some examples in Table 1.

Table 1. Examples of patterns

Biological Patterns	PH Transformations	Descriptions
		This pattern model the activation of the component b by the component a.
		This pattern model the inhibition of the component b by the component a
		This pattern model either the activation of the component c by the component a or the inhibition of the component c by the component b

Estimating the parameters for the PH-simulation from time-series gene expression data. Since the simulation of the execution of the PH actions is done stochastically, we need to relate each action with temporal and stochastic parameters introduced into the PH framework to achieve dynamic refinement [11]. To fire an action in the PH framework we need to provide two parameters: (1) the rate $r = t^{-1}$, where t is the mean time of firing an action, and (2) the stochasticity absorption factor sa, which is introduce to control the variance of firing an action.

For the model components which have a measurement in the time-series data we estimate and integrate r and sa parameters in the PH model. The rest of the components will be assigned default parameters. In order to estimate r_i and sa_i for each action $h_i \in \mathcal{H}$, we need to know the different times t_i when the action could be fired as illustrated in Fig. 5. Each t_i represents the time in which we assume that a component moves from one process to another. Therefore the action that leads this change must be played with rate $r_i = \frac{1}{t_i - t_{i-1}}$. The integer sa represents the window of firing the action at rate r : the larger the sa, the small the variance around r is.

Discretization of times-series data Because the output of a PH simulation are discrete traces of PH components, we discretized continuous experimental data to facilitate the comparison with simulation outputs. When looking at the time-series data (see Fig. 3) one can distinguish a high level of activity in early hours [$0h$-$5h$] and a low level in late

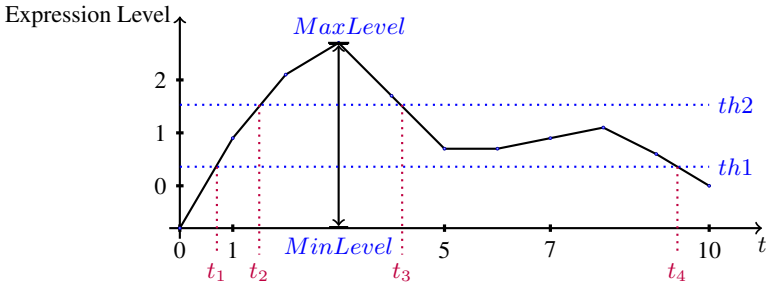


Fig. 5. Estimation of temporal parameters from time series data: The mean firing time of an action that makes a component (gene) change of sub-state is estimated as $r_i = \frac{1}{t_i - t_{i-1}}$. $MaxLevel$ represents the maximum expression of a gene, while $MinLevel$, its minimum expression. The thresholds $th1$ and $th2$ define the PH discrete sub-states (e.g. 0,1,2) of a component according to its gene expression data.

hours [$5h$-$10h$]. We implemented a discretization method to capture these two activation times. For each time-series, we introduced two thresholds *th1* and *th2* (see Fig. 5): *th1* $= \frac{1}{3}(MaxLevel - MinLevel)$ and *th2* $= \frac{2}{3}(MaxLevel - MinLevel)$. In this way, the expression level in the range $[0 - th1]$ is at level 0, the one in the range $[th1 - th2]$ is at level 1, and the one in the last range is at level 2.

2.4 Simulation

We set the same initial conditions to PH components belonging to the same network layer, chosen from the RSTC structure. These initial conditions are detailed below and summarized in Fig. 6.

- **Receptor layer: E-cadherin.** We choose the pulse signal for the input node E-cadherin to be active for a duration of 5 units times in average. We made this choice to take into account the average time of the Calcium stimuli effect.
- **Signaling layer: signaling proteins.** The components in this layer are activated and inhibited with the same rate and the same stochasticity absorption factor. The actions between a controller component A and a controlled component B are constrained so that B is first activated by A and then inhibited. That is, the time interval in which an inhibition action from A to B fires is greater than the time interval in which an activation action from A to B fires. Additionally, these two time intervals must not overlap. These constraints can be seen as reachability constraints from the entry node (E-cadherin) to the output nodes (genes). The values of these parameters are selected by considering the delay of signal transduction from the entry node to the output nodes.
- **Transcription layer: transcription factors.** In this layer, the activation/inhibition over a transcription factor (TF) comes from signaling proteins; however, for all TFs we introduced an auto-inhibition action that represents their degration over time.
- **Genes.** The genes are activited or inhibited according to the estimated values from time-series data.

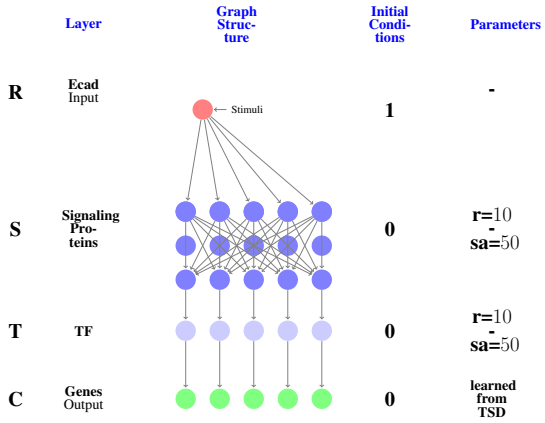


Fig. 6. RSTC network structure and initial conditions assigned to each node in the layer

3 Results

3.1 Automatic generation of PH model from the PID network

PH models are written using the PINT[3] format. PINT implements stochastic simulations and static analyses for computing dynamical properties on very large-scale PH models. For the PINT code generation we used two procedures. The first procedure detects motifs (controller and controlled components) in graphs from the Pathway Interaction Database; the second, generates the PINT code by choosing an adequate concurrency rule, based on synchronization sorts, to represent the motif dynamic in PH. With this method it is possible to convert the whole content of the PID-NCI database into a PH model, as well as individual pre-selected pathways, as is the case for the system under study. It is implemented in Java and available upon request.

3.2 Simulation of Calcium stimulated biological system

We simulated the model with and without the inclusion of the synchronization sort. In the following, we present the results of the simulation.

Without the introduction of the synchronization sort. One can notice in Fig. 7 the occurence of oscillations. Whereas it is not the expected behavior from the biological system, it is coherent with the choice of the modeling and the way the simulator works as explained in Section 2.2. In this simulation cooperation sorts were used to model multiple controllers of a common controlled (target) component. It is important to notice that the intensity of the oscillation is linked with the size of the concurrence, i.e. the number of controllers a controlled component has. Despite the presence of the oscillations, the model reproduces expected dynamical behavoirs namely the dynamic

[3] `http://process.hitting.free.fr`

of components, the signal transduction and takes into account the stochastic and time aspect of the model.

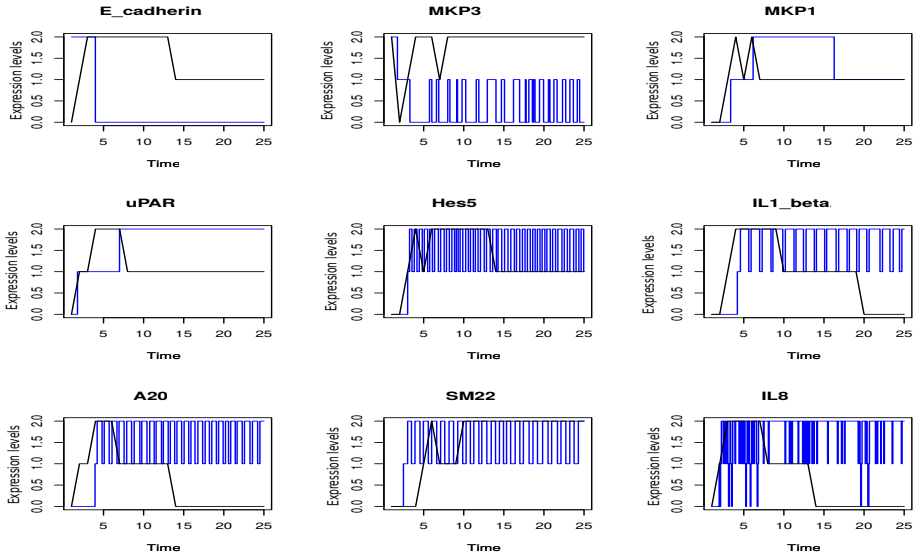


Fig. 7. Results of system simulations without introducing the synchronization sort for 9 genes. The traces representing the discretized time-series data are shown as black lines. The traces representing the simulated traces are shown as blue lines.

With the introduction of the synchronization sort. In Fig. 8 we can see that the introduction of the synchronization sort significatively reduces the impact of concurrency. The result shows a clear elimination of the previously observed oscillation (Fig. 7). Comparing the simulated results with the ones observed experimentally we found four different cases. We found one simulation trace (IL8) that matches the sequence of all the component expression levels perfectly. In this case delays exist among simulation and experiment but these delays are not comparable since experimental time-points are measured in hours and simulation-units for the simulated PH model. We found 4 simulation traces (uPAR, ET1, TfR, SM22) that matched the sequence of experimental discrete expression levels missing one expression-level. We found 6 components (MKP3, MKP1, Hes5, IL1-beta, TNF-alpha, DKK1) in which at least 2 expression levels are missed. Finally we found 1 component (A20) for which we did not reproduce its experimentally observed dynamics.

Simulating biological processes. To validate our model we studied the prediction of non-observed components of such a system and we focused on biological processes

linked to Calcium stimulation, such as keratinocyte-differentiation, cell-adhesion and cell-cycle arrest. Our results are shown in Fig. 9 and confirmed litterature experimental evidence on these processes. In the case of keratinocyte-differentiation, this was a functional behavior measured on the cultured cells upon Calcium stimulation, so there was experimental evidence of this effect before measuring the gene expression. In the case of cell-cycle arrest, the switch-on of this component represents the fact that the E-cadherin stimulated model predicts the stop of growth, as confirmed by litterature in human and mouse keratinocytes [5]. Finally, the cell-adhesion component is predicted to switch-on, also in according to published evidence [14] in human and mouse keratinocytes.

4 Conclusion

This work describes the steps towards the integration of time-series data in large-scale cell-based models. We proposed an automatic method to build a timed and stochastic PH model from pathways of biochemical reactions present in the Pathway Interaction Database (PID). As a case-study we built a model combining signaling and transcription events relevant to keratinocyte differentiation induced by Calcium, which linked E-cadherin nodes and 12 genes, which expression profiles was measured upon Calcium stimulation over time. The interaction graph represented by the model had 293 and 375 edges. We proposed a method to discretize time-series gene expression data, so they can be integrated to the PH simulations and logically explained by the PH stochastic analyses. Additionally, we implemented a method to automatically estimate the temporal and stochastic parameters for the PH simulation, so this estimation process will not be biased by over fitting.

Our results show that we can observe dynamic effects on 11 out of 12 genes, for which 5 of them represent accurate predictions, and 6 missed few dynamic levels. This error may be also a result from the incompleteness of the regulatory information in PID. Moreover, when observing the predicted behavior of biological processes linked to Calcium stimulation, our predictions agreed with experimental and litterature-based evidences. Overall, with this work we show the feasability of modeling and simulating large-scale networks with very few parameter estimation and having good quality predictions.

As perspectives of this work we intend to study the effects of computing automatically the concurrent rules on this system. Also, we intend to improve the model prediction quality by empirically obtaining the dynamics of the system components by performing large stochastic simulations, as well as by implementing static analysis of quantitative properties by adding probabilistic features to the PH static solver.

5 Acknowledgements

This work was supported by a PhD grant from the CNRS and the French region *Les Pays de la Loire* and grants from the German Ministry for Research and Education (BMBF) funding program MedSys (grant number FKZ0315401A) and AGENET (FKZ0315898).

References

1. Grégory Batt, Delphine Ropers, Hidde De Jong, Johannes Geiselmann, Radu Mateescu, Michel Page, and Dominique Schneider. Validation of qualitative models of genetic regulatory networks by model checking: Analysis of the nutritional stress response in escherichia coli. *Bioinformatics*, 21(suppl 1):i19–i28, 2005.
2. Carito Guziolowski, Aristotelis Kittas, Florian Dittmann, and Niels Grabe. Automatic generation of causal networks linking growth factor stimuli to functional cell state changes. *FEBS Journal*, 279(18):3462–3474, 2012.
3. Carito Guziolowski, Santiago Videla, Federica Eduati, Sven Thiele, Thomas Cokelaer, Anne Siegel, and Julio Saez-Rodriguez. Exhaustively characterizing feasible logic models of a signaling network using answer set programming. *Bioinformatics*, 29(18):2320–2326, 2013.
4. Monika Heiner, David Gilbert, and Robin Donaldson. Petri nets for systems and synthetic biology. In *Formal methods for computational systems biology*, pages 215–264. Springer, 2008.
5. C. Kolly, M. M. Suter, and E. J. Muller. Proliferation, cell cycle exit, and onset of terminal differentiation in cultured keratinocytes: pre-programmed pathways in control of C-Myc and Notch1 prevail over extracellular calcium signals. *J. Invest. Dermatol.*, 124(5):1014–1025, May 2005.
6. Aidan MacNamara, Camille Terfve, David Henriques, Beatriz Peñalver Bernabé, and Julio Saez-Rodriguez. State–time spectrum of signal transduction logic models. *Physical Biology*, 9(4):045003, 2012.
7. Mylène Maurin, Morgan Magnin, and Olivier Roux. Modeling of genetic regulatory network in stochastic π-calculus. In *Bioinformatics and Computational Biology*, pages 282–294. Springer, 2009.
8. Alexander Mitsos, Ioannis N Melas, Paraskeuas Siminelakis, Aikaterini D Chairakaki, Julio Saez-Rodriguez, and Leonidas G Alexopoulos. Identifying drug effects via pathway alterations using an integer linear programming optimization formulation on phosphoproteomic data. *PLoS computational biology*, 5(12):e1000591, 2009.
9. Mohammad Mobashir, Burkhart Schraven, and Tilo Beyer. Simulated evolution of signal transduction networks. *PloS one*, 7(12):e50905, 2012.
10. Michael K. Molloy. Performance analysis using stochastic petri nets. *Computers, IEEE Transactions on*, 100(9):913–917, 1982.
11. Loïc Paulevé, Morgan Magnin, and Olivier Roux. Refining dynamics of gene regulatory networks in a stochastic π-calculus framework. In *Transactions on Computational Systems Biology XIII*, pages 171–191. Springer, 2011.
12. Corrado Priami. Stochastic π-calculus. *The Computer Journal*, 38(7):578–589, 1995.
13. Carl F Schaefer, Kira Anthony, Shiva Krupa, Jeffrey Buchoff, Matthew Day, Timo Hannay, and Kenneth H Buetow. Pid: the pathway interaction database. *Nucleic acids research*, 37(suppl 1):D674–D679, 2009.
14. C. L. Tu, W. Chang, and D. D. Bikle. The calcium-sensing receptor-dependent regulation of cell-cell adhesion and keratinocyte differentiation requires Rho and filamin A. *J. Invest. Dermatol.*, 131(5):1119–1128, May 2011.
15. John J Tyson, Katherine C Chen, and Bela Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current opinion in cell biology*, 15(2):221–231, 2003.
16. Namhee Yu, Jihae Seo, Kyoohyoung Rho, Yeongjun Jang, Jinah Park, Wan Kyu Kim, and Sanghyuk Lee. hipathdb: a human-integrated pathway database with facile visualization. *Nucleic acids research*, 40(D1):D797–D802, 2012.

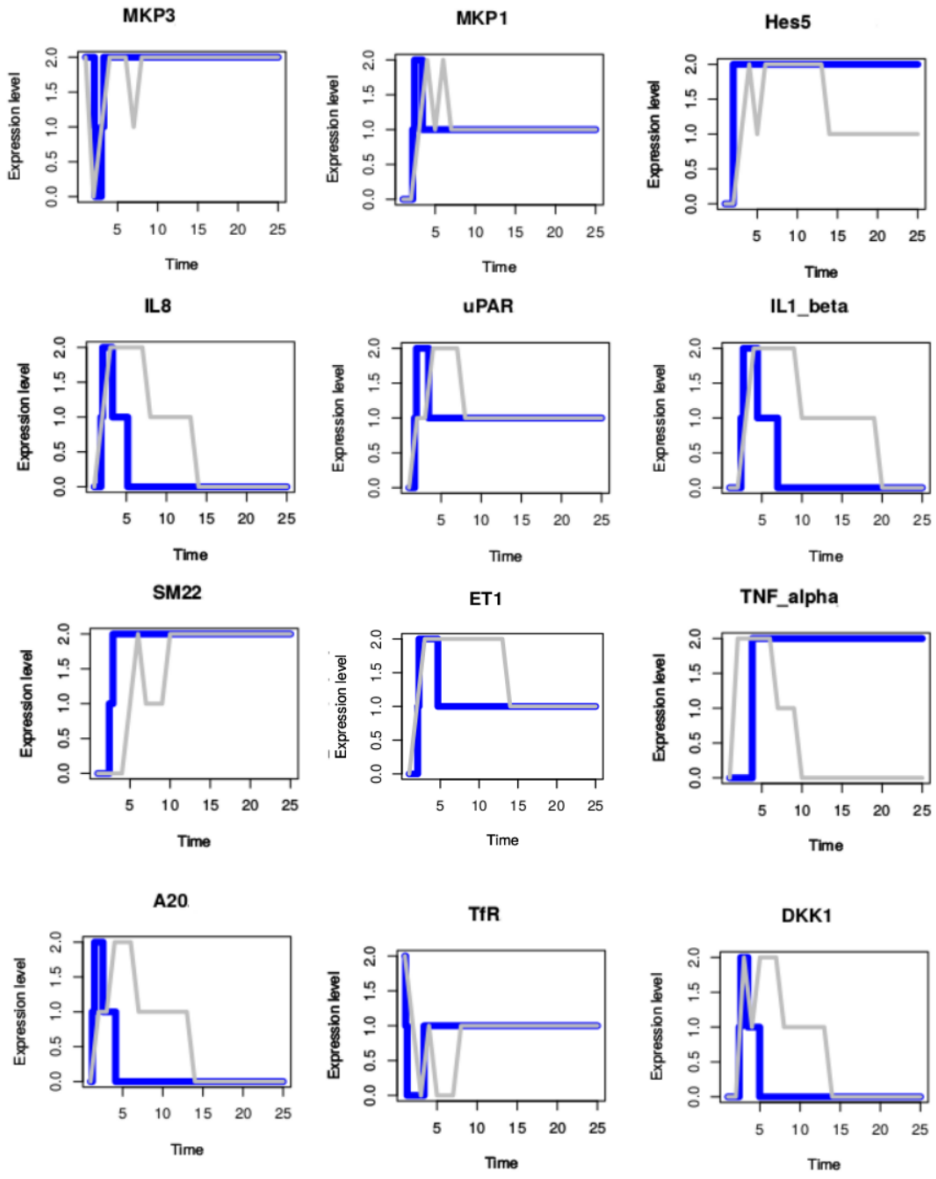


Fig. 8. Results of simulations by introducing the synchronization sort. The gray traces represent the experimental expected behaviors from the discretization of the time-series data. The blue traces show the simulated behavior.

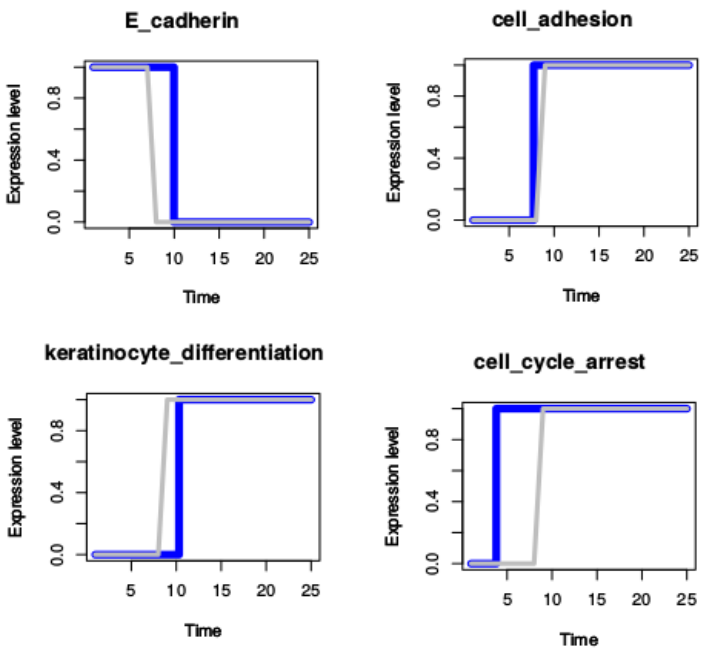


Fig. 9. Results of the prediction of biological processes. The gray traces represent the experimental and litterature-based evidence. The blue traces show the simulated behavior of E-cadherin and three biological processes.