Simulating FAS-induced apoptosis by using P systems*

Smitha Cheruku[1], Andrei Păun[1]**, Francisco J. Romero-Campero[2], Mario J. Pérez-Jiménez[2] and Oscar H. Ibarra[3]

(1. Department of Computer Science/IfM, Louisiana Tech University, P.O. Box 10348, Ruston, LA 71272, USA; 2. Department of Computer Science and Artificial Intelligence, University of Seville, Avda. Reina Mercedes s/n 41012, Sevilla, Spain; 3. Department of Computer Science, University of California -Santa Barbara, Santa Barbara, CA 93106, USA)

Abstract In contrast to differential equations, P systems are an unconventional model of computation which takes into consideration the discrete character of the quantity of components and the inherent randomness that exists in biological phenomena. The key feature of P systems is their compartmentalised structure which represents the heterogeneity of the structural organisation of the cells, and where one can take into account the role played by membranes in the functioning of the system, for example signalling at the cell surface, selective uptake of substances from the media, diffusion across different compartments, etc.

We show here that P systems can be a reliable tool for Systems Biology and could even outperform in some cases the current simulation techniques based on differential equations. We will also use a strategy based on the well known Gillespie algorithm but running on more than one compartment called Multi-compartmental Gillespie Algorithm.

Keywords: FAS-induced apoptosis, P systems, deterministic waiting times algorithm, multi-compartmental Gillespie algorithm.

Understanding biosignalling pathways is essential for designing effective therapeutic approaches to several important diseases. For example, the FAS-induced apoptotic signalling pathway was shown to be one of the most relevant processes for understanding and combating cancer, AIDS and neurodegenerative diseases such as Parkinson's disease, Alzheimer, etc. With several pathways unraveled in the last years, each one with its own unique structure and complexity, there is an increasing need to model these signalling cascades due to their complex nature. Because there is usually immense data collected for only one pathway, it is almost always hard to understand the pathway without the help of computer simulators. For better understanding of the FAS-induced apoptosis we are proposing a new way (actually two different approaches) of simulating the pathway by using P systems.

A typical group of biosignalling pathways is known to lead to apoptosis (also known as programmed cell death). Apoptosis is a mechanism which helps the unwanted, injured, or improperly developed cells to commit suicide playing a fundamental role in the fight of the organism against cancers. Thus, aberrations in apoptotic responses to death signals contribute to cancer development, resistance to treatment, but also the reverse problem: autoimmune diseases. One major mechanism for inducing the apoptosis is through the activation of death receptors. Among the death receptors the signalling pathways for FAS-induced apoptosis are best characterised at the moment. We believe that computational modelling can play an important role to advance understanding the complex signalling behavior of this pathway.

Modelling FAS-induced apoptosis (or any biosignalling pathway) can be done in many ways, the traditional approach is at the moment the use of differential equations. It is argued that the use of differential equations is not the best approach for simulating processes that involve low number of molecules/objects as the ordinary differential equations (ODEs) are assuming large populations of molecules and are modelling the changes in the concentration/numbers of molecules of a particular species. For low numbers of molecules it is questionable if ODEs provide an accurate modelling[1].

A model, an abstraction of the real-world onto a mathematical/computational domain, highlights some key features while ignoring others that are assumed to be not relevant. A good model should have at least four properties: relevance, computability, understandability and extensibility[2]. A model must be rel-

* A. Păun gratefully acknowledges the support in part by LA BoR RSC grant LEQSF (2004-07)-RD-A-23 and NSF Grants IMR-0414903 and CCF-0523572; O. H. Ibarra gratefully acknowledges the support in part by NSF Grants CCR-0208595, CCF-0430945 and CCF-0524136

** To whom correspondence should be addressed. E-mail: apaun@latech.edu

evant capturing the essential properties of the phenomenon investigated; and computable so it can allow the simulation of its dynamic behavior, as well as the qualitative and quantitative reasoning about its properties. An understandable model will correspond well to the informal concepts and ideas of molecular biology. Finally, a good model should be extensible to higher levels of organisations, like tissues, organs, organism, etc., in which molecular systems play a key role. We believe that P systems possess all these properties.

The approach based on differential equations is usually referred to as macroscopic chemistry since they model the average evolution of the concentration of chemical substances across the whole system.

The microscopic approach considers the molecular dynamics for each single molecule involved in the system taking into account their positions, momenta of atoms, etc. This approach is computationally intractable because of the number of atoms involved, the time scale and the uncertainty in many of the cellular components.

Our approach is referred to as mesoscopic chemistry[3]. Like in the microscopic approach one considers individual molecules like proteins, DNA and mRNA but ignores many other molecules like water and non-regulated parts of the cellular machinery. Besides this, the position and momenta of the molecules are also not modelled, instead one deals with the statistics of which reactions occur and how often. This approach is more tractable than microscopic chemistry but it provides a finer and better understanding than the macroscopic chemistry.

Another observation is that the P system paradigm focuses on the compartmental structure that is exhibited by the cells; in each compartment one has different rules and objects, and the system moves from one configuration to the next one by obeying the rules and using only the objects available in each compartment. These are the features that we want to simulate in the signalling pathways, in this respect the use of P systems appears to be a natural approach.

1 P Systems using deterministic waiting times algorithm

The P systems used for modelling FAS-induced apoptosis consider the quantity of components in the cascade to be discrete. We also have that the individual chemical reactions are asynchronous and occur discretely at different intervals of time. We reach this extension of P systems by considering that the rules can take various amounts of time (not only 1 time cycle as they were defined in the original variants). We also have the rules in the system obeying the Law of Mass Action: the reaction rate depends proportionally on the product of the concentrations of the reactants.

An incipient model of P systems of this type was reported last year[4], where we have defined a system in which molecules can associate and dissociate (the current rules are of this form) and also rules can have different "application times", in the sense that different rules can take different amounts of time to finish. The current model extends that model further by modelling the Law of Mass Action in the sense that the same rule r could take 5 clock cycles to finish in a particular configuration of the system, or 10 clock cycles in another configuration of the system. The difference is in the number of the reactants for the reaction; this number is influencing the actual speed with which the reaction is applied (i.e. in the second configuration we would have approximately half of the number of reactants with respect to the first configuration).

In this paper we work with a variant of P systems being a tuple

$$\Pi = (O, L, \mu, M_1, M_2, \cdots, M_n, R_1, \cdots, R_n),$$

where O is a finite alphabet of symbols representing objects (proteins and complexes of proteins); L is a finite alphabet of symbols representing labels for the compartments (membranes); μ is a membrane structure containing $n \geqslant 1$ membranes labelled with elements from L; $M_i = (w_i, l_i)$, $1 \leqslant i \leqslant n$, are pairs which represent the initial configuration of membrane i: $l_i \in L$ is its label, and $w_i \in O^*$ is the initial multiset; R_i, $1 \leqslant i \leqslant n$, are finite sets of rules associated with the membrane i which are of the form $u[v]_{l_i} \rightarrow u'[v']_{l_i}$, where $u, v, u', v' \in O^*$ are finite multisets of objects and l_i is the label of membrane i.

Next, we discuss in more detail the rules that we will use in this paper, to model protein—protein interactions taking place in the compartmentalised structure of the living cell.

Transformation, complex formation and dissociation rules:

$$[a]_l \rightarrow [b]_l \quad (1)$$

$$[a + b]_l \rightarrow [c]_l \quad (2)$$

$$[a]_l \rightarrow [b + c]_l \quad (3)$$

where $a, b, c \in \Sigma$ and $l \in L$. These rules are used to specify chemical reactions taking place inside a compartment of type $l \in L$, more specifically they represent the transformation of a into b; the formation of a complex c from the interaction of a and b; and the dissociation of a complex a into b and c.

Diffusing in and out:

$$[a]_l \rightarrow a[]_l \quad (4)$$

$$a[]_l \rightarrow [a]_l \quad (5)$$

where $a \in \Sigma$ and $l \in L$. When chemical substances move or diffuse freely from one compartment to another we use these types of rules, where a moves from or to a com-partment of type l.

Binding and unbinding rules:

$$a[b]_l \rightarrow [c]_l \quad (6)$$

$$[a]_l \rightarrow b[c]_l \quad (7)$$

where $a, b, c \in \Sigma$ and $l \in L$. Using rules of the first type we can specify reactions consisting in the binding of a ligand swimming in one compartment to a receptor placed on the membrane surface of another compartment. The reverse reaction, unbinding of substance from a receptor, can be described using the second rule.

Recruitment and releasing rules:

$$a[b]_l \rightarrow c[]_l \quad (8)$$

$$c[]_l \rightarrow a[b]_l \quad (9)$$

where $a, b, c \in \Sigma$ and $l \in L$. With these rules we represent the interaction between two chemicals in different compartments whereby one of them is recruited from its compartment by a chemical from the other compartment, and then the new complex remains in the latter compartment. In a releasing rule a complex, c, located in one compartment can dissociate into a and b, remaining a in the same compartment as c, and b being released into the other compartment.

Each reaction r has associated a rate constant c_r computed using those in [5] according to [6]. The time needed for the reaction r to finish will then be computed in the following way: $wt_r = \frac{1}{c_r * |A| * |B|}$, where $|A|$ and $|B|$ represent the number of molecules of the two reactants A, B of the reaction r. On the other hand, if the reaction simulated is a dissociation or other first order reaction in A, the reaction time is computed as $wt_r = \frac{1}{c_r * |A|}$.

We give the algorithm as follows:

① Calculate the WT for all the reactions in all the membranes, call it "waiting time"—T. Select the reaction, r, with the shortest WT.

② Update the waiting time for the rest of the reactions by subtracting T and add it to the simulation time ($t = t + T$).

③ Apply r only once, thereby updating the multiplicities of respective symbols.

④ Recalculate the waiting times only for those reactions which are in the compartments affected by the applied rule. For each such reaction compare the new waiting time with the existing WT and keep the shortest one.

⑤ Sort all reactions with respect to their waiting times and again take the reaction with the lowest waiting time.

⑥ Repeat steps 2 to 5 until the desired simulating time is reached.

We believe that this strategy represents the most natural way of defining the evolution of the P system. It is based on the fact that the rate of reaction is directly proportional to the product of concentrations of reactants. We calculate the velocity of reactions using this fact and, likewise the waiting time.

We implemented in Java a simulator for P systems with these dynamics. It accepts as input a Systems Biology Markup Language (SBML) file containing the rules to be simulated and initial number of molecules in the system. We used Cell Designer package to generate the SBML source file for the reactions.

2 Results obtained for modelling FAS-induced apoptosis

We have implemented all the rules described in [5] for type I and type II of FAS-induced apoptotic pathway starting with the stimulation of FASL (FAS Ligand) until the activation of the effector Caspase-3. In total we simulated 99 rules in the P system working on 53 proteins and protein complexes. The afore-

mentioned paper also provides the experimental data obtained from the Jurkat cell line, which is a type II pathway dominant cell type. We compared our results with both the experimental data and with the simulation data reported in [5]. The comparison is given in Fig. 1.

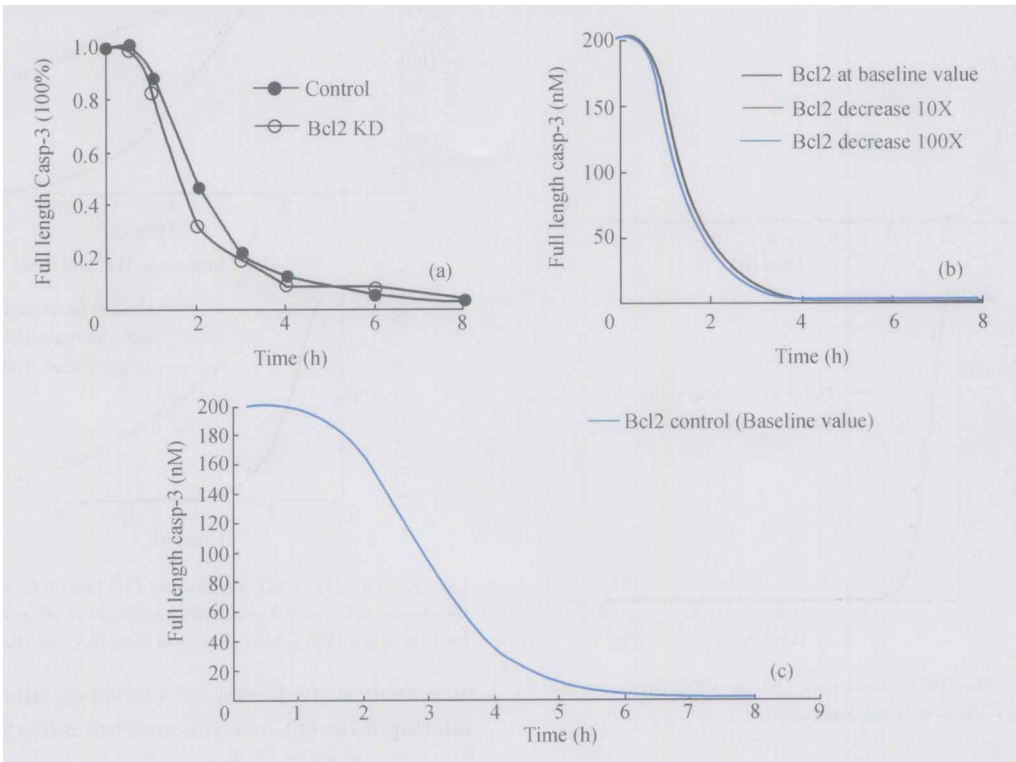

Fig. 1. Comparison between experimental data (a), previous ODE simulation data (b) and the new DWT simulation data (c).

The simulation data from [5] was obtained by simulating the rules using ordinary differential equations (ODE). The ODEs were solved numerically by the Entelos PhysioLab Modeler implementing a Runge-Kutta method, thus calculating the concentration changes for all the molecules in the system.

One of the major proteins in the pathway, Caspase-3 was compared to the experimental data. In the Entelos simulation, Caspase-3 was activated at 4 hours, and it was considered close to the experimental results where it was obtained that it was activated at 6 hours (see Fig. 2).

The same pathway is modelled in the membrane computing framework using the same reactions and initial conditions. The Caspase-3 activation dynamics is studied when Bcl2 is at baseline value. Caspase-3 is activated in our simulator after about 7 hours which is a very good approximation of the experimental data and also the simulated pathway in [5].

We pass now to perform all the simulation runs reported in [5]: several changes in the initial concentrations were considered such as increases in the initial concentration for Bcl2 of tenfold and one hundredfold, also increase in Caspase-8 of twentyfold.

These simulations were run to see the sensitivity of the overall pathway to changes in the initial concentrations of the aforementioned proteins. Another change considered in [5] was to check whether Bcl2 should be defined as binding to Bax, Bid, tBid (truncated Bid) or both Bax and tBid. Several simulations with the respective changes in the set of reactions to model each assumption were considered in the original paper and we matched the same simulations.

2.1 Modelling the type I pathway behavior by modifying the model

There are cells which are not sensitive to Bcl2 over expression as described in [7]. In these cells Caspase-3 is activated through the type I pathway, independent of the type II pathway. Scaffidi et al.[7] have suggested that the type of pathway is chosen based on the concentration of Caspase-8 generated in active form following FASL binding. If the Caspase-8 concentration is high, then the Caspase-3 is activated directly. On the other hand, if the concentration of

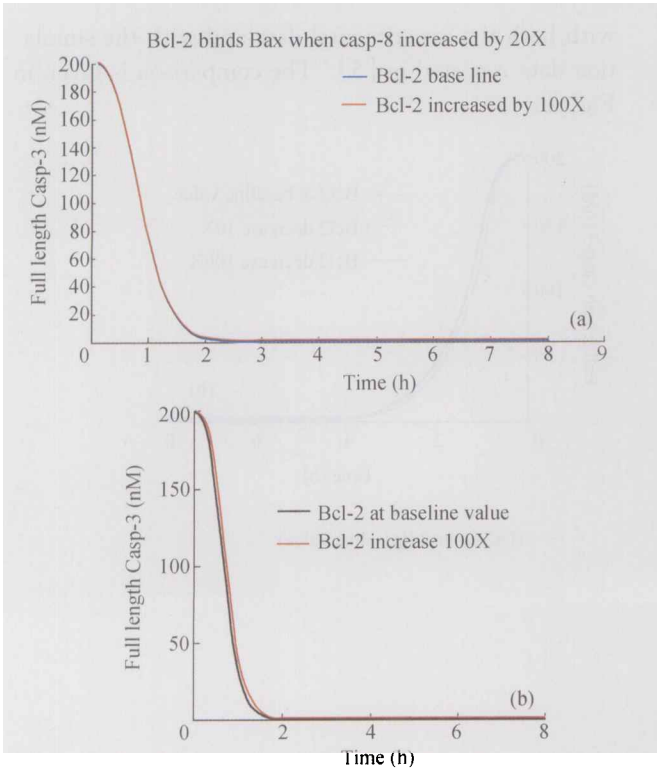

Fig. 2. (a) The DWT simulation; (b) the ODE simulation for the change in Caspase-8 initial concentration.

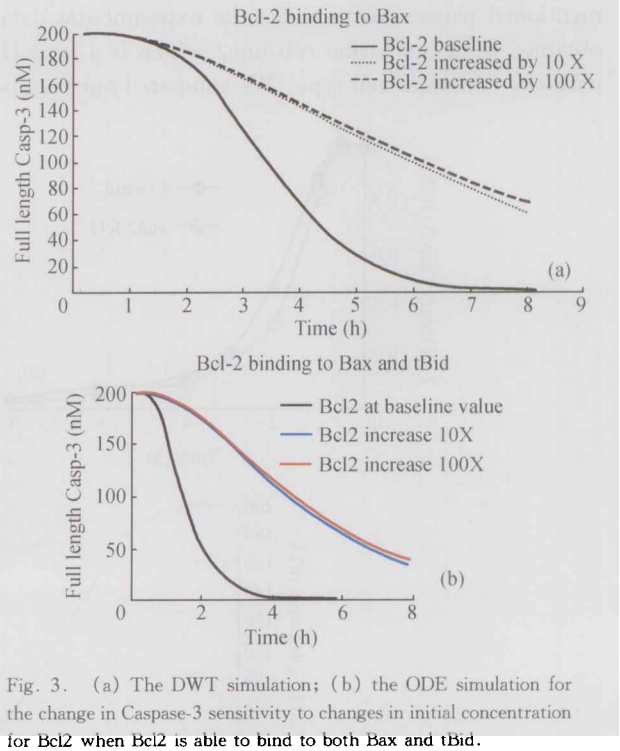

Fig. 3. (a) The DWT simulation; (b) the ODE simulation for the change in Caspase-3 sensitivity to changes in initial concentration for Bcl2 when Bcl2 is able to bind to both Bax and tBid.

Caspase-8 is low, the type II pathway is chosen so that the system is amplifying the death signal through the mitochondria to be able to induce the cell death. To test this hypothesis, the active Caspase-8 formation is increased by having the initial concentration of Caspase-8 set to a value 20 times greater than its baseline value while everything else was kept the same in the system. We performed the same simulation with the increase in Caspase-8 initial concentration, which resulted in faster Caspase-3 activation also in our simulation and is in good agreement with the previous paper.

The Bcl2 concentration is also increased 100 times to test the sensitivity of Caspase-3 activation to Bcl2. Fig. 3 shows that the Caspase-3 activation is not sensitive to the increase in Bcl2 concentration, which is the hallmark for the type I pathway dominant behavior.

The consistency between the framework and the experimental results in [5] validate the new simulation model.

2.2 Bcl2's effect on type II pathway

Bcl2 is an inhibitor of the type II pathway. We now analyse the Caspase-3 activation kinetics by considering different mechanisms through which Bcl2 can block the type II pathway. In [8], [9], or [10] the authors suggest that Bcl2 might bind with (a) Bax, (b) Bid, (c) tBid, or (d) bind to both Bax and tBid to block the mitochondrial pathway. We created four different models of P systems having the rules $r_1, \cdots, r_{96}, r_{97}$ for case (a), $r_1, \cdots, r_{95}, r'_{96}, r'_{97}$ for case (b), $r_1, \cdots, r_{95}, r''_{96}, r''_{97}$ for case (c) or the rules $r_1, \cdots, r_{97}, r_{98}, r_{99}$ for case (d). All the other rules remain the same for all the cases. We refer the interested reader to [11] for details of the model. The dynamics of Caspase-3 activation is studied by varying the Bcl2 concentration 10 times or 100 times the baseline value. It was concluded that Bcl2 binding to both Bax and tBid is the most efficient mechanism for the pathway in comparison with the results obtained for cases (a), (b) or (c). The same conclusions were obtained also after using our simulator for all the previous changes in the pathway. Due to space limitations, we do not include here all the results of the simulations, but only case (d) as a comparison between the ODE simulator and our simulator.

We will now describe the same simulation of the FAS-induced apoptosis by using a strategy that is based on the well known Gillespie's algorithm but running on more than one compartment and so it will

be called Multi-compartmental Gillespie Algorithm[12].

3 Modelling FAS-induced apoptosis by the multi-compartmental gillespie algorithm

Gillespie's algorithm[6] provides an exact method for the stochastic simulation of systems of bio-chemical reactions; the validity of the method is rigorously proved and it has been already successfully used to simulate various biochemical processes[13]. Moreover, the Gillespie's algorithm is used in the implementation of stochastic π-calculus, and in its application to the modelling of biological systems[14]. An extension of the classical Gillespie's algorithm called Multi-compartmental Gillespie Algorithm was introduced in [12]. This method is developed by taking into account the fact that, with respect to the original algorithm where only one volume is studied, in P systems we have a membrane structure delimiting different regions or compartments, each one can be seen as a volume with its own set of rules, besides the application of a rule inside a compartment can also affect the content of another one; for example the application of a communication rule.

P systems assume rates that determine the speed of reactions. In this respect, we associate to each rule a stochastic constant which represents the average number of application of the rule per time unit. This stochastic constant will be used to compute the probability of applying a rule in a given configuration[12]. This is necessary to characterise the reality of the phenomenon to be modelled.

4 Some results and discussion

We used a simulator designed in Scilab available from [11] and using the multi-compartmental Gillespie algorithm.

The goal of this signalling cascade is to translate the environmental FASL concentration into the cytoplasmic phosphorylated Caspase-3 concentration.

Briefly the signal transduction takes place as follows. FASL is the ligand of a transmembrane receptor, FAS. When FASL binds to FAS a conformational change takes place in the receptor producing the complex FASC. The cytoplasmic domain of this complex recruits FADD (FAS-associated death domain), actually is able to recruit three molecules of FADD (a trimer). Once FADD is bound to the complex FASC, either Caspase-8 or FLIP can be recruited competitively. The binding of FLIP prevents Caspase-8 to bind to FASC and thus it prevents Caspase-8 phosphory-lation; in this sense FLIP acts as a repressor in the signalling cascade. Once at least two molecules of Caspase-8 have been recruited, a dimer Caspase-8_2^{P41} is released into the cytoplasm and there it is phosphorylated to produce Caspase-8*. In the cytoplasm a series of protein-protein interactions take place, through the two different pathways discussed before, that have as final goal the phos-phorylation of Caspase-3.

In what follows we present some graphs, depicting the evolution over time of the concentration or number of molecules of some of the important proteins in the pathway.

In Fig. 4, it is showed the evolution over time of the concentration of Caspase-3. Note that at the beginning there is a very slow phosphorylation of Caspase-3 and suddenly almost all the molecules get phosphorylated.

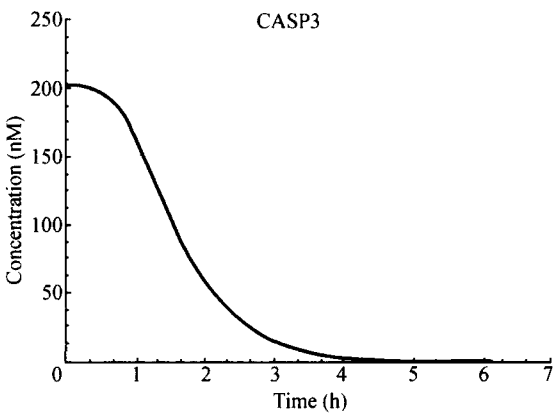

Fig. 4. CASP3 activation dynamics over time.

In the evolution over time of the concentration and number of molecules of Caspase-8 (see Fig. 5), it can be observed that Caspase-8 is consumed rapidly, in a few minutes all the molecules that can be recruited by the number of activated receptors are consumed. On the other hand, the FASC-$FADD_2$-Caspase-8_2 protein complex exhibits a sudden increase in the number of molecules and then in about one hour the number of the molecules arrives in a sub-20 range where it stays throughout the simulation.

Note in Fig. 6 (a) that the complex FASC-FADD is a very instable molecule. When these molecules are formed they rapidly recruit other

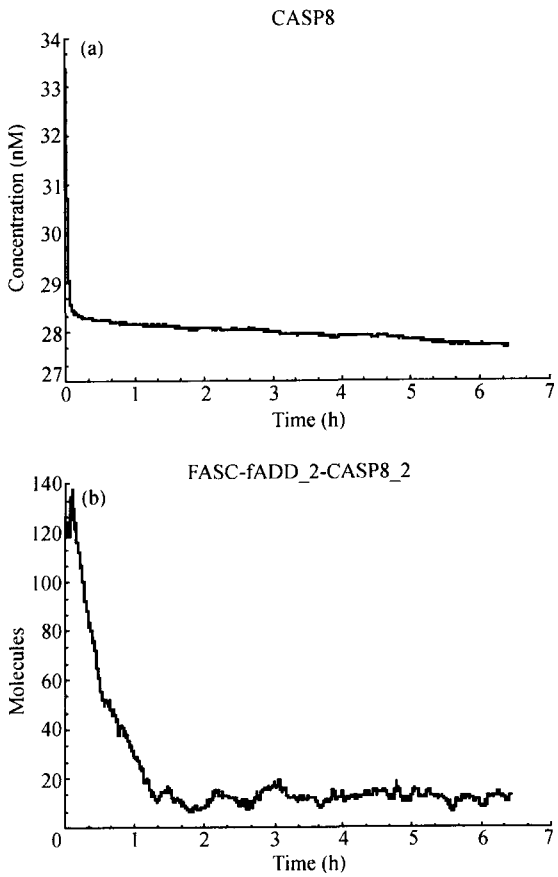

Fig. 5. Evolution over time of the concentration of CASP8 (a) and number of molecules of FASC-FADD$_2$-CASP8$_2$(b).

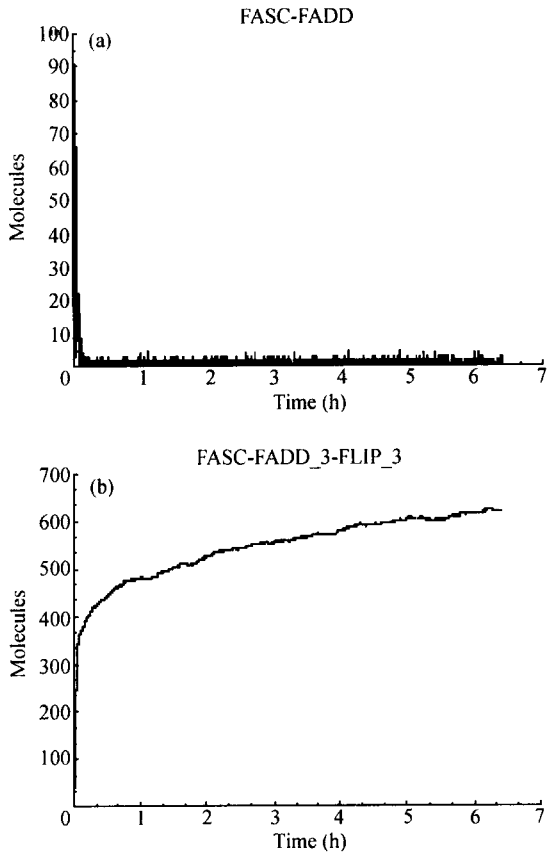

Fig. 6. Evolution over time of the number of molecules of FASC-FADD (a) and FASC-FADD$_3$-FLIP$_3$(b).

molecules (see Fig. 6 (b)), and therefore only a few molecules are shown in the evolution of the system.

5 Conclusion

We have simulated the FAS-induced apoptosis using the P systems, and DWT rule application or multi compartmental Gillespie algorithm. Both simulations showed good correlation with the experimental data reported in the literature. Since the simulators are rule based, they prove flexible, and can easily integrate any future discoveries/updates in the pathway. The simulator gives 8 hours worth of signalling pathway time in about one and one half hour, and it seems to be very fast when compared to similar simulators based on ODE's.

We plan to implement the algorithm in C + + and make freely available the source-code to the community. We also believe that improvements in running time can be achieved by using data structures techniques (implementing a heap for the finding of the next reaction to be considered) which could reduce the run time of the algorithm from $O(n^2 \log n)$ to $O(n \log n)$. Introducing a nondeterministic behavior for the DWT could also be proved interesting. More comparisons between the three methods for simulating the signalling pathways are also planned.

References

1 Firth CAJM and Bray D. Stochastic simulation of cell signalling pathways. In: Computational Modelling of Genetic and Biochemical Networks, 2001

2 Regev A and Shapiro E. The π-calculus as an abstraction for biomolecular systems: In: (eds. Gabriel Ciobanu and Grzegorz Rozenberg). Modelling in Molecular Biology, Springer Berlin, 2004

3 Van Kampen NG. Stochastic Processes in Physics and Chemistry. Elsevier Science B. V., Amsterdam, The Netherlands, 1992

4 Ibarra OH and P$\breve{a}$un A. Counting time in computing with cells. In: Proceedings of DNA Based Computing, DNA11, 2001, London, Ontario, 25—36

5 Hua F, Cornejo M, Cardone M, et al. Effects of Bcl-2 levels on FAS signalling-induced caspase-3 activation: molecular genetic tests of computational model predictions. The Journal of Immunology, 2005, 175(2): 985—995 and correction 175(9): 6235—6237

6 Gillespie DT. Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry, 1977, 81(25): 2340—2361

7 Scaffidi C, Fulda S, Srinivasan A, et al. Two CD95 (APO-1/Fas) signalling pathways. The Embo Journal, 1998, 17: 1675—1687
8 Wang K, Yin XM, Chao DT, et al. BID: a novel BH3 domain-only death agonist. Genes & Development, 1996, 10: 2859—2869
9 Oltavi ZN, Milliman CL and Korsmeyer SJ. Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. Cell, 1993, 74(4): 609—619
10 Cheng EH, Wei MC, Weiler S, et al. BCL-2, BCL-XL sequester BH3 domain-only molecules preventing BAX-and BAK-mediated mitochondrial apoptosis. Molecular Cell, 2001, 8: 705—711
11 P System Simulator: http://www. dcs. shef. ac. uk/~ marian/PSimulatorWeb/PSystemMF. htm
12 Pérez-Jiménez MJ and Romero-Campero FJ. P systems, a new computationl modelling tool for systems biology. Transactions on Computational Systems Biology VI, LNBI, 2006, 4220: 176—197
13 Meng TC, Somani S and Dhar P. Modelling and simulation of biological systems with stochasticity. Silico Biology, 2004, 4(3): 293—309
14 Priami C, Regev A, Shapiro E, et al. Application of a stochastic name-passing calculus to representation and simulation of molecular processes. Information Processing Letters, 2001, 80: 25—31