On Modeling Signal Transduction Networks

Alberto Castellini, Giuditta Franco

Verona University, Computer Science Dept.
Strada Le Grazie 15, 37134 Verona, Italy
E-mails: {alberto.castellini, giuditta.franco}@univr.it

Summary. Signal transduction networks are very complex processes employed by the living cell to suitably react to environmental stimuli. Qualitative and quantitative computational models play an increasingly important role in the representation of these networks and in the search of new insights about these phenomena. In this work we analyze some graph-based models used to discover qualitative properties of such networks. In turn, we show that MP systems can naturally extend these graph-based models by adding some qualitative elements. The case study of integrins activation during the lymphocyte recruitment, a crucial phenomenon in inflammatory processes, is described, and a first MP graph for this network is designed. Finally, we discuss some open problems related to the qualitative modeling of signaling networks.

1 Introduction

Biological signal transduction is a series of processes employed by the living cell to convert signals coming from the external environment [17, 1]. Signal conversions usually involve sequences of chemical reactions among proteins which generate complex networks. By these mechanisms the cell can suitably react in order to accomplish its biological functions.

The discovery of hidden interaction mechanisms supports the development of new medical treatments and drugs for specific diseases, thus, in the last decades many efforts have been addressed to "decipher" the interactions among the actors of signaling networks. Despite of the great medical and pharmaceutical interest for these networks, a lot of them are still completely or partially unknown, because of their huge size and high complexity. So far, the best results have been achieved at a qualitative level, by the representation and the analysis of protein interactions [18]. Such results have been mainly reached by means of suitable data structures (mainly graphs) that support the topological analysis of biological networks, while an open challenge concerns the modeling of dynamics related to the protein activation. This could provide significant improvements for the prediction of signaling network behaviors.

Both *qualitative* and *quantitative* modeling of metabolic and signal transduction networks have drawn large advantages from the use of computational models, which disclose new features of these systems by processing a great deal of data [33]. The most used mathematical models for the dynamical analysis of biological networks are the traditional systems of ordinary differential equations (ODE). They represent a biological system by a set of mathematical equations, where every equation rules the temporal evolution of a substance. Unfortunately, ODE models have some drawbacks, such as the hard definition of equations from observed phenomena, since a deep knowledge about microscopic molecular kinetics is required [20]. Thus, several discrete and bio-inspired models have been proposed, in order to symbolically describe cellular processes.

P systems, or *membrane systems*, were introduced in [28] as a new computational model which takes its inspiration from the structure and functioning of the living cell. This model is rooted in the context of formal language theory and it is based on *multisets* and *membranes* rewriting for which many computational universality results have been achieved [29]. This mathematical framework has been often employed to model biological processes underlying metabolism [2, 5, 27], pathologies [14, 13] and signaling [12, 26, 32].

Metabolic P systems, shortly MP systems, were recently introduced in [24] and developed, along with different lines, see for example [6, 21, 22, 23]. The aim of this non-conventional mathematical framework is properly the modeling of metabolic dynamics. In fact the development of this model has been based on the *mass partition principle*, which defines the transformation rate of object populations, according to a suitable generalization of chemical laws [20].

Some equivalence results have been proved in [11] and in [9] between MP systems and, respectively, autonomous ODE and Hybrid Functional Petri nets. The dynamics of several biological processes has been effectively modeled by means of MP systems [3], among them: the Belousov-Zhabotinsky reaction (in the Brusselator formulation) [5, 6], the Lotka-Volterra dynamics [5, 24], the SIR (Susceptible-Infected-Recovered) epidemic [5, 3], the leukocyte selective recruitment in the immune response [5], the Protein Kinase C activation [6], the circadian rhythms [3], the mitotic cycles in early amphibian embryos [23, 12], a Pseudomonas quorum sensing model [7] and the *lac* operon gene regulatory mechanism in glycolytic pathway [9].

In this work we investigate the possibility to employ MP systems for modeling signal transduction networks. We point out that some chemical laws which regulate metabolic processes cannot be applied to signaling processes, and several measurement problems can arise, due to current experimental limitations. With respect to models of metabolism, the focus has to be moved from substance transformations (complex creation and disintegration) to protein activation (e.g., phosphorylation and dephosphorylation).

In section 2 it is presented a qualitative model currently used to represent protein interactions [18], and some motivations and problems related to quantitative models are discussed. In Section 3 it is proposed the case study of a signaling

network involved in the lymphocyte recruitment, while in the last section some approaches are addressed for modeling the above network by MP systems.

2 From Qualitative to Quantitative Modeling

Several biological systems can be represented, at a *qualitative* level, by a network of elements connected by functional interactions. Some examples are the nervous system (physically connected neurons), the immune system (where cells and molecules are connected with different kinds of interaction) and also signaling systems (physically connected intracellular molecules). These networks can be symbolically represented by *graphs*, mathematical structures used to model pairwise relations (*edges*) between objects (*nodes*) from a certain collection. In signal transduction networks, nodes act as molecules, typically proteins, and arcs represent the capability of a molecule to activate or deactivate another molecule (Figure 1).

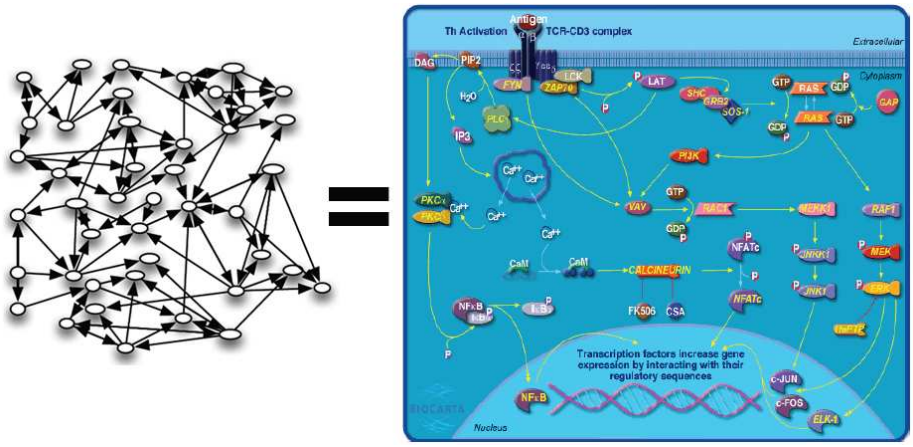

Fig. 1. A qualitative representation of a signal transduction network by a directed graph. Nodes correspond to molecules and arcs represent activation or deactivation [18].

Graph theory investigates structural properties of graphs, and these properties can assume a biological meaning if the graph represents a specific biological network. Topological analyses based on graph theory often address interesting questions about single elements of the network, clusters of elements, or the entire network. There exist specific computational tools [31] to visualize networks and analyze topological properties, such as the node degree (the number of edges connected to a node), the average degree of the network, the shortest path between two nodes, the average shortest path of the network, the network diameter,

clustering coefficients, the presence of network motifs (e.g., loops) or network statistical properties [18]. In a signal transduction network, the knowledge of such properties most of the times allows the identification of substances which regulate the process.

The qualitative approach described above is valid for a statical analysis of networks. While structural properties can be found, the temporal evolution of a network instead cannot be "simulated" to forecast dynamical behaviors (e.g., oscillations) occurring under different conditions. *Quantitative* models aim at extending qualitative models to achieve this target. They rely on new experimental methods which support the measurement of real systems (e.g., activation levels or concentration values), absolutely necessary to validate models. Namely, in the last decade, the wide improvements of high-throughput data acquisition techniques in molecular biology have made possible to screen and to analyze the expression of entire genomes, as well as to assess large numbers of proteins and to characterize in detail the metabolic state of a cell population, although several problems must be still overcome. On the other hand, new mathematical and computational techniques have been conceived to infer coherent theories and models from huge amounts of experimental data [33].

Systems of Ordinary Differential Equations (ODE) have been largely used for the quantitative modeling of signal transduction networks, but lately some network-oriented and bio-inspired models are overcoming several drawbacks of traditional ODE models. They allow a new insight about biological processes, which cannot be obtained using the "glasses" of classical mathematics [2].

In the next section we focus on a cellular process of great immunological interest, with the goal to design a model for this case study. It is a partially unknown network obtained by long and complex laboratory experiments and proposed very recently in [8].

3 A Case Study

Inflammatory processes in living organisms activate a tissue-specific recruitment of leukocytes, that relies on the complex functional interplay between some surface molecules of leukocytes circulating in the blood and the endothelial cells covering the blood vessel. Leukocyte recruitment into tissues requires extravasation from blood, by a process of transendothelial migration. Three major steps have been identified in the process of leukocyte extravasation: tethering-rolling of free-flowing white bloods cells, leukocyte activation and their arrest by the adherence to endothelial cells. After the arrest *diapedesis* happens, namely leukocytes pass from blood to the tissue beyond endothelial cells [30].

The recruitment process takes place when molecules called *chemokines* are produced by the epithelium and by bacteria that have activated the inflammation. Chemokines bind with *receptors* located on the leukocyte membrane, then activating an internal signaling network. The main output of this network is the

activation of *integrins*, different receptors that interact with endothelial *counter-receptors* slowing down the leukocyte speed, until its arrest. If we call A the initial state, with leukocytes quickly circulating into the blood, B the state of rolling, C the state of activation, and D the state of adhesion, three main phases can be recognized: $A \rightarrow B$ ruled by receptor-receptor interactions, $B \rightarrow C$ ruled by chemokine-receptor interactions, and $C \rightarrow D$ ruled by integrin-receptor interactions. This process was modeled by P systems in [14], where the concept of receptor was integrated with objects transformed and moved through membranes by rewriting systems having rules with priority.

Very recent works have discovered a minimal signaling module activated by chemokines and controlling the integrin activation during the whole process of recruitment of *lymphocytes* B and T, two specific types of leukocyte [8]. Figure 2 shows the entire module (in the center) surrounded by the lymphocyte membrane in which receptors *CXCR4* and integrins *LFA-1* are placed. Elliptical nodes represent types of molecules, *PA* and *PIP2* are second messengers, continuous arrows indicate experimentally demonstrated direct activations (with physical interactions and complex formations), dashed arrows indicate indirect effects, flat line endings indicate inhibitions, empty circle endings indicate second messenger production, the arch ending indicates a direct binding, that is without any enzymatic activation.

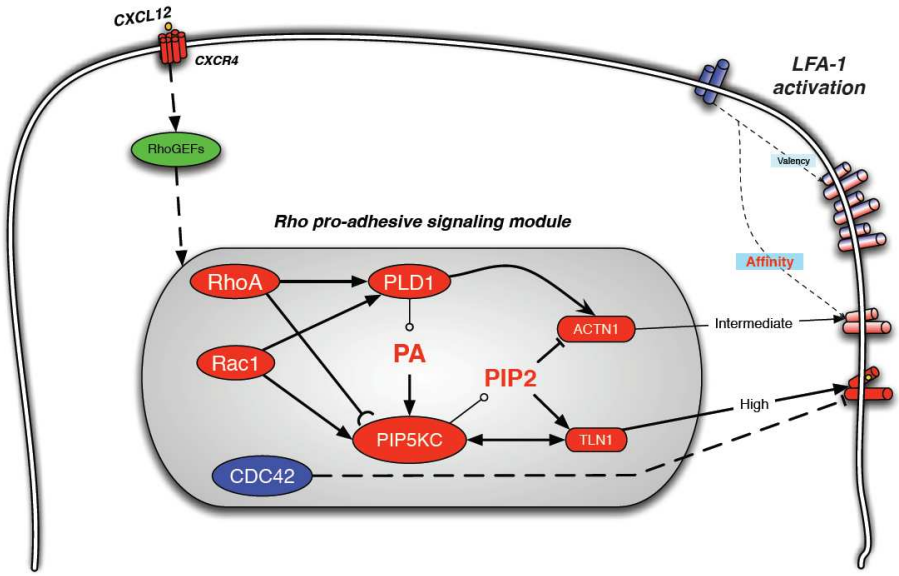

Fig. 2. Model of the Rho-signaling module, activated by chemokines and controlling conformer-specific *LFA-1* affinity triggering during the lymphocyte homing [8].

The *Rho pro-adhesive signaling module* proposed in [8] and reported in Figure 2 takes its input from the receptor *CXCR4*, which is activated by chemokines *CXCL12* and indirectly activates *RhoGEFs* molecules . In turn, RhoGEFs activate the three small GTPase proteins *RhoA*, *Rac1* and *CDC42* inside of the functional module. RhoA and Rac1 activate PLD1, thus leading to phosphatidic acid (*PA*) accumulation. *PIP5KC* bind to RhoA and it may be activated by Rac1 and PA. At the same time, PLD1 may interact with alpha-actinin1 (*ACTN1*), facilitating the interaction between *ACTN1* and the integrins *LFA-1* leading their transition to intermediate affinity state. LFA-1 transitions among low, intermediate and high level are depicted from the top to the bottom (they are all integrins) on the right side of Figure 2. Simultaneously, activated PIP5KC triggers the local accumulation of phosphatidylinositol 1-4-5 phosphate (*PIP2*), which has a central role in the transition of LFA-1 from intermediate to high affinity state, and for the leukocyte firm arrest. The increase of the PIP2 concentration, in fact, may inhibit ACTN1, facilitating its detachment from LFA-1, and may activate Talin1 (*TLN1*), driving the final transition to the high affinity state. We also notice that PIP5KC may activate directly TLN1, and this may promote direct transition of LFA-1 to the high affinity state. In this complex context, *CDC42* plays a "negative" regulatory role by preventing LFA-1 activation.

The signaling network reported in Figure 2 is a qualitative representation of the integrin LFA-1 activation during the lymphocyte homing. An interesting open problem from the biomedical viewpoint is the discovery of qualitative parameters which rule the dynamics of the network over time, such as activation speeds or delays. The development of a qualitative model would allow notable predictions of system behaviors in presence of normal or abnormal (e.g., pathological and pharmacological) conditions. However, qualitative models could imply quite a few experimental problems, because only some quantities are measurable over time, while most of them cannot be measured by current technologies. Our final goal is the discovery of regulative mechanisms of the Rho pro-adesive signaling module, namely starting from the curve of LFA-1 affinity in lymphocytes stimulated by chemokines, showed in Figure 3.

4 Signaling Networks Modeled by MP Systems

As we hinted in the previous sections, the main difficulties of signaling modeling, compared to metabolic modeling, are (i) the lack of stoichiometric coefficients that rule the ratios of chemical interactions, (ii) the consequent non-applicability of some chemical laws, such as the mass conservation law, the Avogadro principle and the Dalton principle, that constrain metabolic models, (iii) the lack of data acquisition techniques to measure activation and concentration levels over time. These difficulties could make inappropriate several methods yet employed for modeling metabolic pathways. New meaningful models must be identified to predict signaling network behaviors, hopefully despite the current lack of information about the underlying phenomena.

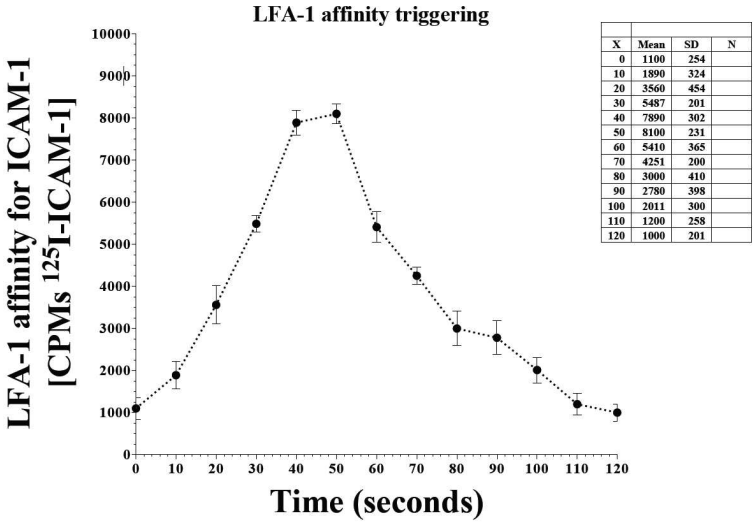

Fig. 3. LFA-1 affinity of a cluster of lymphocytes stimulated by chemokines.

In the following we propose a representation of the network of Figure 2 by means of MP systems, since it is a network-oriented quantitative model which has several features in common with the graph-based qualitative models described in Section 2. *Petri nets* have been also employed to "graphically" model signaling networks from a qualitative point of view [25], in fact their equivalence with MP systems has been proved in [9].

MP systems are deterministic P systems developed to model the dynamics of biological phenomena related to metabolism. They naturally extend graph-based models because they consist of (i) a set of **substances** X, each one associated to (concentration or activation) quantity when observed, (ii)a set of **reactions** R, that move substances, (iii) a set of **parameters** V (such as pressure, temperature,...) each equipped with an evolution function, and (iv)a set Φ of **flux regulation maps**, which state the amount of substances consumed/produced by every reaction in any system transition. The dynamics δ of this model is represented by the evolution of substances and parameters in every temporal interval τ starting from the initial state σ_0.

A graphical representation of MP systems by means of *MP graphs* has been introduced in [23]. MP graphs depict biochemical reactions as bipartite graphs with two levels, in which the first level describes the *stoichiometry* of reactions, while the second level expresses the *regulation* which tunes the flux of every reaction (i.e. the quantity of chemicals transformed at each step) depending on the state of the system (see for example Figure 4).

Recent works aim at deducing MP models from suitable macroscopic observations of given metabolic behaviors along a certain number of steps [21, 19]. In order to assist biologists in the simulation of MP systems we implemented a Java

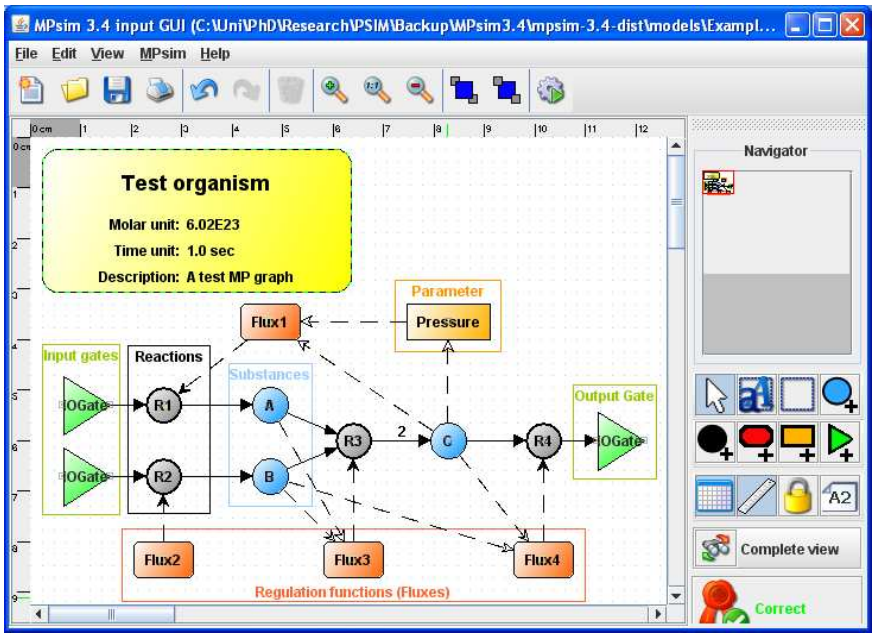

Fig. 4. An example of MP graph visualized by the MPsim 3 graphical user interface. Frame labels point out MP system elements in the MP graph representation.

tool called *MPsim* [4]. The current release of the software employs a friendly user interface to define MP models by means of MP graphs [23], and it produces the plotting of curves given by the system dynamics. The last developments of the software involve a new plugin framework for solving parameter estimation, analysis, visualization and importation tasks in order to increase the system capabilities.

Figure 5 shows an MP graph representation of the signal transduction network depicted in Figure 2. For every chemical involved in the network we set a *substance node* (ellipses) and a linked *reaction node* (circles), which update the substance quantity at each step. We model the quantities associated to PA and PIP2 as concentrations, and the quantities associated to the other substances as activation levels. The *activation level* of a protein expresses its capability to activate other chemicals in the signaling cascade, therefore it is a very important feature for qualitative modeling. Instead, second messengers such as PA and PIP2 tune their interaction with other chemicals depending on their concentration.

Every reaction node is linked to a *flux node* (squares) which computes the quantity added or removed by the reaction at each step, depending on the system state. Dashed arcs from a substance node x to a flux node φ_y (which regulates the substance y) indicates that x is a "regulator" of y. For instance, the activation arc from *RhoA* to *PLD1* in Figure 2 is mapped to the "regulation" arc from the substance node *RhoA* to the flux node *F6*, because the updating of *PLD1*

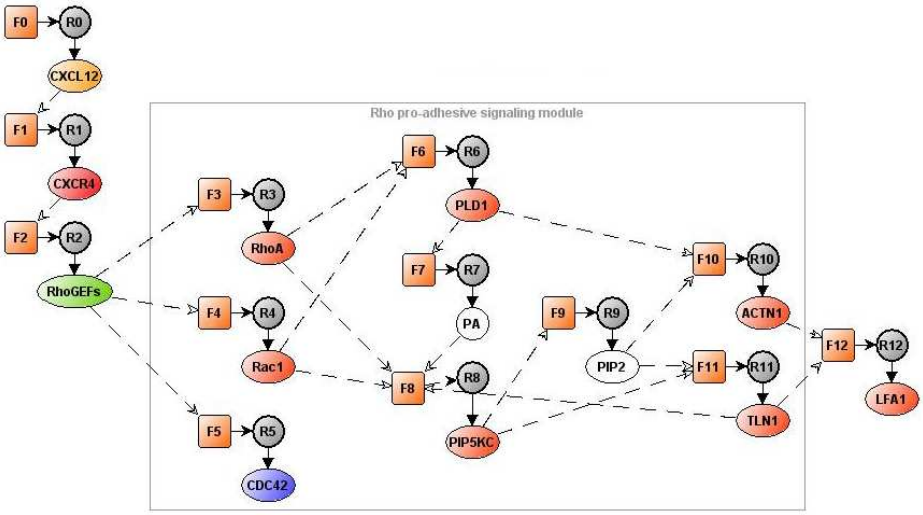

Fig. 5. An MP graph representation of Rho-signaling module activated by chemokines and controlling conformer-specific *LFA-1* affinity.

activation level over time depends on the *RhoA* activation level. All the arcs of Figure 2 have been mapped to MP graph dashed arcs in this way.

The mapping of a qualitative model to an MP graph seems quite a simple task though a crucial problem keeps open: the definition of significant activation speeds and delays which suitably fit the observed dynamics of the system (Figure 3). These parameters affect the regulation functions Φ of the MP system and they would allow good predictions of the system behaviors.

As a future research, we intend to attack this problem by bio-inspired methodologies, like neural networks, and evolutionary techniques, like genetic programming, already proposed by John R. Koza in [16] for the automatic synthesis of metabolic pathways and genetic networks. The latter computational technique could suggest solutions even for the topology validation of qualitative models. Indeed, activation arcs of qualitative models (as that in Figure 2) are drawn when an experiment proves the interaction between two molecules. However, the topology we have of a network is possibly not complete, because some interactions have not been discovered yet. In these cases quantitative computational models, which are based on wide data observations, can point out interaction lacks and suggest new experiments to discover different networks.

References

1. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular biology of the cell.* Garland Science, New York, 4th edition, 2002.

2. F. Bernardini and V. Manca. Dynamical aspects of P systems. *BioSystems*, 70:85–93, 2003.
3. L. Bianco. *Membrane models of biological systems.* PhD thesis, University of Verona, 2007.
4. L. Bianco and A. Castellini. Psim: a computational platform for Metabolic P systems. In G. Eleftherakis, P. Kefalas, G. Păun, G. Rozenberg, and A. Salomaa, editors, *8th International Workshop on Membrane Computing, WMC 2007, Thessaloniki, Greece, June 25-28, 2007, Revised, Selected and Invited Papers*, volume 4860 of *Lecture Notes in Computer Science*, pages 1–20. Springer, 2007.
5. L. Bianco, F. Fontana, G. Franco, and V. Manca. P systems for biological dynamics. In Ciobanu et al. [10], pages 81–126.
6. L. Bianco, F. Fontana, and V. Manca. P systems with reaction maps. *International Journal of Foundations of Computer Science*, 17(1):27–48, 2006.
7. L. Bianco, D. Pescini, P. Siepmann, N. Krasnogor, F. J. Romero-Campero, and M. Gheorghe. Towards a P systems pseudomonas quorum sensing model. In Hoogeboom et al. [15], pages 197–214.
8. M. Bolomini-Vittori, A. Montresor, C. Giagulli, D. Staunton, B. Rossi, M. Martinello, G. Constantin, and C. Laudanna. Conformer-specific LFA-1 activation and turnover regulation by chemokine-triggered rho signaling module in human lymphocytes. *Nature Immunology*, 2008. Submitted.
9. A. Castellini, G. Franco, and V. Manca. Toward a representation of Hybrid Functional Petri Nets by MP systems. In Y. Suzuki, G. Păun, A. Adamatzky, M. Hagiyra, and H. Umeo, editors, *Natural Computing, 2nd International Workshop on Natural Computing, IWNC 2007, Nagoya University, Japan, December 10-12, 2007*, pages 1–12, 2007.
10. G. Ciobanu, G. Păun, and M. J. Pérez-Jiménez, editors. *Applications of Membrane Computing.* Natural Computing Series. Springer, 2006.
11. F. Fontana and V. Manca. Discrete solutions to differential equations by metabolic P systems. *Theoretical Computer Science*, 372(2-3):165–182, 2007.
12. G. Franco, P. H. Guzzi, V. Manca, and T. Mazza. Mitotic oscillators as MP graphs. In Hoogeboom et al. [15], pages 382–394.
13. G. Franco, N. Jonoska, B. Osborn, and A. Plaas. Knee joint injury and repair modeled by membrane systems. *Biosystems*, 91(3), March.
14. G. Franco and V. Manca. A membrane system for the leukocyte selective recruitment. In *4th International Workshop on Membrane Computing, WMC 2003, Tarragona, Spain, July 17-22, 2003, Revised Papers*, 2004.
15. H. J. Hoogeboom, G. Păun, G. Rozenberg, and A. Salomaa, editors. *Membrane Computing, 7th International Workshop on Membrane Computing, WMC 2006, Leiden, Netherlands, July 17-21, 2006, Revised, Selected, and Invited Papers*, volume 4361 of *Lecture Notes in Computer Science.* Springer, 2006.
16. J.R. Koza, M.A. Keane, M.J. Streeter, W. Mydlowec, J. Yu, and G. Lanza. *Genetic Programming IV : Routine Human-Competitive Machine Intelligence (Genetic Programming).* Springer, 2003.
17. G. Krauss. *Biochemistry of signal transduction and regulation.* Wiley-VCH, 2nd english edition, 2001.
18. C. Laudanna. Introduzione all'analisi topologica delle reti biologiche. Url: http://www.cbmc.it.
19. V. Manca. Log-gain principles for Metabolic P systems. In *G. Rozenberg Festschrift.* To appear.

20. V. Manca. The Metabolic Algorithm: Principles and applications. *Theoretical Computer Science*. To appear.
21. V. Manca. Discrete simulations of biochemical dynamics. In M. H. Garzon and H. Yan, editors, *13th International Meeting on DNA Computing, DNA13, Memphis, TN, USA, June 4-8, 2007, Revised, Selected Papers*, volume 4848 of *Lecture Notes in Computer Science*, pages 231–235. Springer, 2007.
22. V. Manca. Metabolic P systems for biochemical dynamics. *Progress in Natural Sciences,* Invited Paper, 17(4):384–391, 2007.
23. V. Manca and L. Bianco. Biological networks in metabolic P systems. *BioSystems*, 91(3):489–498, March 2008.
24. V. Manca, L. Bianco, and F. Fontana. Evolutions and oscillations of P systems: Applications to biochemical phenomena. In G. Mauri, G. Păun, M. J. Pérez-Jiménez, G. Rozenberg, and A. Salomaa, editors, *5th International Workshop, WMC 2004, Milan, Italy, June 14-16, 2004, Revised Selected and Invited Papers*, volume 3365 of *Lecture Notes in Computer Science*, pages 63–84. Springer, 2005.
25. H. Matsuno, Y. Tanaka, H. Aoshima, A. Doi, M. Matsui, and S. Miyano. Biopathways representation and simulation on Hybrid Functional Petri Nets. *In Silico Biology*, 3(3):389–404, 2003.
26. A. Paun, M. J. Pérez-Jiménez, and F. J. Romero-Campero. Modeling signal transduction using P systems. In Hoogeboom et al. [15], pages 100–122.
27. M. J. Pérez-Jiménez and F. J. Romero-Campero. P systems: a new computational modelling tool for systems biology. *Transactions on Computational Systems Biology VI, Lecture Notes in Bioinformatics, 4220*, pages 176–197, 2006.
28. G. Păun. Computing with membranes. *Journal of Computer and System Sciences*, 61(1):108–143, 2000.
29. G. Păun. *Membrane Computing. An Introduction.* Springer, Berlin, Germany, 2002.
30. L. A. Segel and I. R. Cohen (eds). *Design principles for the Immune System and Other Distributed Autonomous Systems.* Oxford University Press, 2001.
31. The Cytoscape Web Site. Url: http://www.cytoscape.org/.
32. Y. Suzuki and H. Tanaka. Modeling p53 signaling pathways by using multiset processing. In Ciobanu et al. [10], pages 203–214.
33. E. Voit, A. R. Neves, and H. Santos. The intricate side of systems biology. *PNAS*, 103(25):9452–9457, June 20 2006.