# Goldbeter's mitotic oscillator entirely modeled by MP systems

Vincenzo Manca and Luca Marchetti

University of Verona, Department of Computer Science
Strada Le Grazie 15, 37134 Verona, Italy.
Tel: +390458027981 Fax: +390458027068
E-mail: vincenzo.manca@univr.it, luca.marchetti@univr.it

**Abstract.** MP systems are a class of P systems introduced for modeling metabolic processes. Here we apply an algorithm, we call Log-Gain Stoichiometric Stepwise Regression (LGSS), to Golbeter's oscillator. In general, LGSS derives MP models from the time series of observed dynamics. In the case of Golbeter's oscillator, we found that by considering different values of the resolution time $\tau$, different analytical forms of regulation maps were appropriate. By means of a suitable MATLAB implementation of LGSS, we automatically generated 700 MP models ($\tau$ varying from $10^{-3}$ min to $700 \cdot 10^{-3}$ min with increments of $10^{-3}$ min). Many of these models exhibit a good approximation, and have second degree polynomials as regulation maps. These results provide an experimental evidence of LGSS adequacy.

## 1 Introduction

Any living organism has to maintain processes which: i) introduce matter of some kinds from the external environment, ii) transform internal matter by changing the molecule distribution of a number of biochemical species (substances, metabolites), and iii) expel matter that is not useful or dangerous to the organism. The molecule distribution identifies the metabolic state of the system in question, and can be represented as a multiset over a set of molecular species $A, B, \ldots, Z$.

An important problem of systems biology is the mathematical definition of a dynamical system that explains the observed dynamics of a phenomenon under investigation, by taking into account what is already known about the phenomenon. When this is possible, then we can hope that a greater knowledge of the phenomenon is gained.

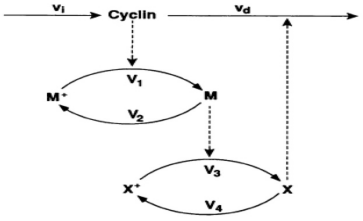
An important line of research of biological modeling is aimed at defining new classes of discrete models avoiding some limitations of classical continuous models based on ordinary differential equations (ODE). In fact, very often, the evaluation of the kinetic reaction rates in differential models is problematic because it may require measurements hardly accessible in living organisms. Moreover, these measurements dramatically alter the context of the investigated processes. In contrast to ODEs, Metabolic P systems (MP systems) [11, 9, 8, 10], based on Păun's P systems [14], were introduced for modeling *metabolic systems* by means of suitable multiset rewriting grammars.

In MP systems no single instantaneous kinetic is addressed, but rather the variation of the whole system under investigation is considered, at discrete time points, separated

by a specified macroscopic interval $\tau$. The dynamics is given along a sequence of steps and, at each step, it is governed by partitioning the matter among reactions which transform it. The log-gain theory of MP systems [8] is aimed at reconstructing the *flux regulation maps* associated to the metabolic transformations. Metabolic P systems proved to be promising in many contexts and their applicability was tested in many situations where differential models are prohibitive due to the unavailability or the unreliability of the kinetic rates [10, 12].

Here we apply an algorithm, we call Log-Gain Stoichiometric Stepwise Regression (LGSS), to Golbeter's oscillator given in Table 1 [3, 4, 5]. In this manner, we generate automatically 700 models of this oscillator, which, for the most part, provide the same order of approximation of Golbeter's model. Moreover, by considering the phenomenon at different time grains, we obtain different models and in many cases the analytical form of these models is simpler than Golbeter's model.

The fundamental mechanism of mitotic oscillations concerns the periodic change in the activation state of a protein produced by the *cdc2* gene in fission yeast or by homologous genes in other eukaryotes. The simplest form of this mechanism is found in early amphibian embryos (see [5] at page 24). Here (see the picture in the left part of Table 1) cyclin (C) is synthesized at a constant rate and triggers the transformation of inactive ($M^+$) into active ($M$) *cdc2* protein, which leads to the formation of a complex known as M-phase promoting factor (*MPF*). MPF triggers mitosis, but at the same time $M$ elicits the activation of a protease from state $X^+$ to $X$. The active protease then degrades cyclin resulting in the inactivation of $cdc2$. This brings the cell back to initial conditions and a new division cycle can take place. The ODE presented in the right part of Table 1 is the differential model of dynamics described in Figure 1, where $C, M, X$ are the concentrations of $C, M, X$ respectively and $1-M, 1-X$ are the concentrations of $M^+, X^+$ respectively (the definitions of the parameters of the ODE model of table 1 are not simple and are not relevant for our further discussion, however they can be found in [3]).




$$\frac{dC}{dt} = v_i - v_d X \frac{C}{K_d+C} - k_d C$$
$$\frac{dM}{dt} = V_{M1} \frac{C}{K_c+C} \frac{(1-M)}{K_1+(1-M)} - V_2 \frac{M}{K_2+M}$$
$$\frac{dX}{dt} = M V_{M3} \frac{(1-X)}{K_3+(1-X)} - V_4 \frac{X}{K_4+X}$$

**Table 1.** Goldbeter's oscillator, which has a cycle of about 25 min [3].

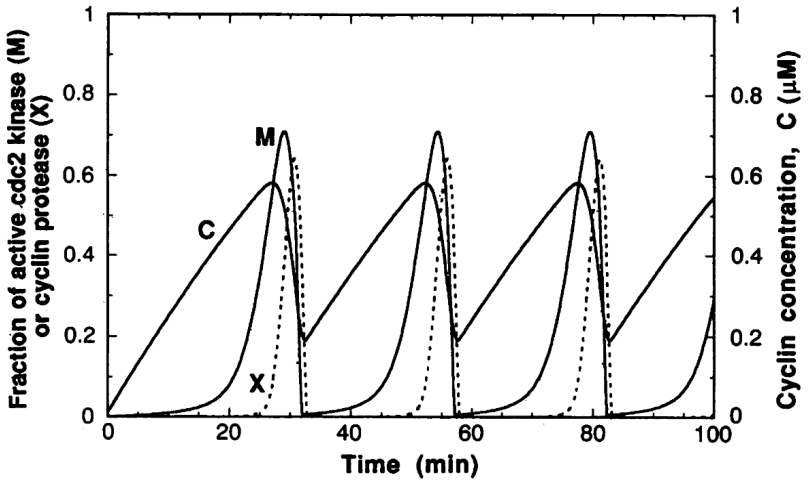



**Fig. 1.** A numerical solution of the set of differential equations (right part of Table 1) comprising the model introduced by A. Goldbeter (figure taken from [3]).

## 2 MP systems

*Metabolic P systems*, or P metabolic systems, represent metabolic processes in a discrete mathematical framework. The letter P of MP systems comes from the theoretical framework of P systems introduced by Gheorghe Păun [14] in the context of membrane computing. In fact MP systems are a special class of P systems introduced in 2004 [11] to express metabolism in a discrete mathematical setting.

A metabolic P system is essentially a multiset grammar where multiset transformations are regulated by functions. Namely, a multiset rule like $A + B \rightarrow C$ means that a number $u$ of molecules of kind $A$ and the same number $u$ of molecules $B$ are replaced by $u$ molecules of type $C$. The value of $u$ is the *flux* of the rule application. Assume to consider a system at some time steps $0, 1, 2, \ldots, t$, and consider a substance $x$ that is produced by rules $r_1, r_3$ and is consumed by rule $r_2$. If $u_1[i], u_2[i], u_3[i]$ are the fluxes of the rules $r_1, r_2, r_3$ respectively, in the passage from step $i$ to step $i + 1$, then the variation of substance $x$ is given by:

$$x[i + 1] - x[i] = u_1[i] - u_2[i] + u_3[i] \tag{1}$$

In a MP system it is assumed that in any state the flux of each rule is provided by a function, called *regulator* of the reaction. Substances, reactions, and regulators (plus parameters which are variables different from substances occurring as arguments of regulators) specify a discrete dynamics at steps indexed in the set $\mathbb{N}$ of natural numbers. Moreover, a *temporal interval* $\tau$, a conventional *mole size* $\nu$, and substances masses are considered, which specify the time and population (discrete) granularities respectively. They are *scale factors* that do not enter directly in the definition of the dynamics of a system, but are essential for interpreting it at a specific physical level of mass and time granularity.

From a mathematical point of view, a MP system M of type $(n, m, k)$, that is, of $n$ substances, $m$ reactions and $k$ parameters (this type will be implicitly assumed), is specified as follows (see also [10]):

**Definition 1 (MP system).** *A MP system is a discrete dynamical system specified by a construct*

$$M = (S, R, H, \Phi, \tau, \nu, \mu)$$

*where $S, R$ are finite disjoint sets, and the following conditions hold, with $n, m, k \in \mathbb{N}$:*

- *$S$ is a set of $n$ substances (the types of molecules) determining, for any metabolic state of the system, a vector $X$ of substance quantities which varies on $\mathbb{R}^n$;*
- *$R$ is a set of $m$ reactions specified by $m$ pairs $(r_1^-, r_1^+), \ldots, (r_m^-, r_m^+) \in \mathbb{N}^n \times \mathbb{N}^n$, composed by the left and right vectors of the reactions (relative to the reactants and to the products respectively). The matrix $\mathbb{A} = (r_1^\#, \ldots, r_m^\#)$ is the stoichiometric matrix associated to the reactions having as columns the stoichiometry balances of the rules;*
- *$H : \mathbb{N} \to \mathbb{R}^k$ is a function providing, at each step $i \in \mathbb{N}$, the vector $H[i]$ of parameters;*
- *$\Phi = (\varphi_1, \ldots, \varphi_m)$ is a vector of regulators (or flux regulation functions), where $\Phi : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$ provides the fluxes of reactions corresponding to any global state of the system, that is, a pair in $\mathbb{R}^n \times \mathbb{R}^k$ constituted by the metabolic state and by the parameter vector. Given a reaction $r$, the substances and the parameters which occur as arguments of the corresponding regulator $\varphi_r$ are called the tuners of the reaction $r$.*
- *$\tau \in \mathbb{R}$ is the time interval between two consecutive steps;*
- *$\nu \in \mathbb{R}$ is the number of molecules which gives a (conventional) mole in the model;*
- *$\mu \in \mathbb{R}^n$ is the vector of the mole masses of substances.*

*Given a vector $X[0] \in \mathbb{R}^n$ relative to an initial state of a given system, the dynamics of $M$ is specified by the following vector recurrent equation, called $EMA[i]$ (Equational Metabolic Algorithm), where $\times$ is the usual matrix product, and, in dependence on the context, $+$ is the usual sum or the component-wise vector sum:*

$$X[i+1] = \mathbb{A} \times U[i] + X[i] \tag{2}$$

*providing the state of the system $X[i+1]$, for each step $i \in \mathbb{N}$, by means of the vector of fluxes $U[i] = (u_r[i] \mid r \in R)$ where $u_r[i] = \varphi_r(a[i], b[i], \ldots)$ and $a[i], b[i], \ldots$ are components of $X[i]$, $H[i]$ which are the tuners of reaction $r$.*

A MP system is completely described by a *MP grammar* where multiset rewriting rules (reactions) are given with the corresponding regulators (plus parameter evolution functions and scale factors for a complete specification of the system). A MP grammar can be also specified by a *MP graph* where the relationships between reactions and regulators appear in a more direct way. An example of MP graph, which represents the Golbeters oscillator [3], is given in Figure 2.

A Java software, called *MetaPlab*, was developed starting from a prototypal version. MetaPlab is downloadable from the official site of MetaPlab software[1]. This platform

[1] `http://mplab.scienze.univr.it`.

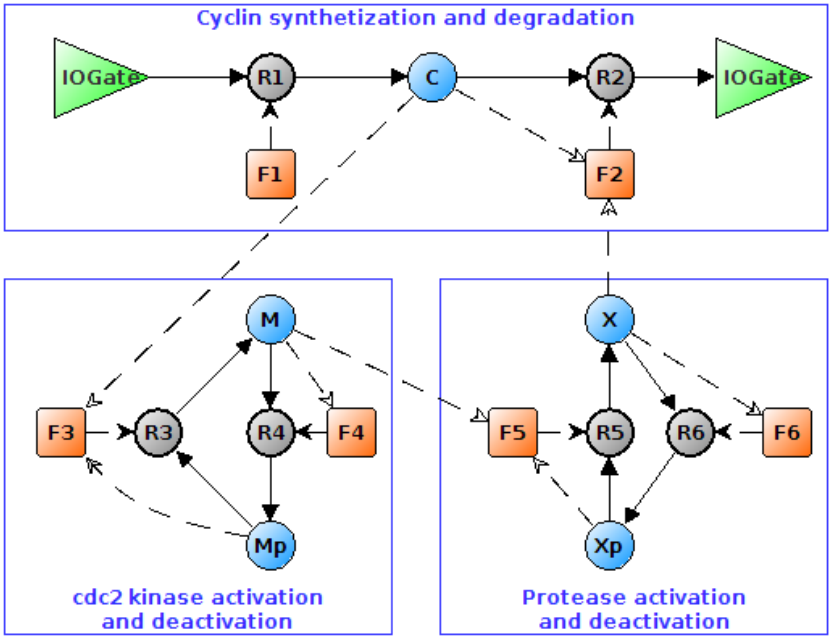



**Fig. 2.** A MP graph. Nodes: triangles represent matter introduction and expulsion, circles $C, M, Mp, X, Xp$ stand for substances, circles $R1, R2, \ldots, R7$ for reactions, rounded corner rectangles for regulators, and rectangles for parameters. Edges: transformation edges go from substances to reactions (consumption) and from reactions to substances (production), regulation edges go from regulators to reactions, and influence edges go from substance or parameters (tuners) to regulators.

enables the user to design MP models by means of some useful graphical tools, to simulate their dynamics, and to automatise some procedures which can help the user to develop new models. MetaPlab is based on an extensible set of plugins, namely Java tools, for solving specific tasks relevant in the framework of MP systems. A guide for this software is available at the official site of MetaPlab software.

### 2.1 The log-gain principle of MP Systems

The log-gain principle was introduced in MP systems theory for solving the following inverse dynamic problem [8, 10]. Given a time series $(X[i], H[i]) \in \mathbb{R}^{n+k}$ (for $i = 0, 1, 2, \ldots t$) of some consecutive states and parameters of a metabolic system (at a time interval $\tau$), is it possible to deduce a corresponding time series of vectors $U[i] \in \mathbb{R}^m$ which put in the equation (2) provide the time series of substance quantities? This is the dynamical problem of reaction flux discovery. The deduction of time series $U[i]$ is related to the time granularity $\tau$ of the systemic logic governing the matter transformations of the observed metabolic states. When vectors $U[i]$ are known, the discovery of maps $\Phi$ which provide $U[i]$, in correspondence to the vectors $(X[i], H[i])$,

is a typical problem of approximation which can be solved with standard techniques of mathematical regression.

An important remark is due in this context. The approach of flux discovery is essentially observational, macroscopic, and global, in a sense which is opposite to the perspective of differential models, which is infinitesimal, and local. In fact, we do not intend to discover the real kinetics responsible for the biochemical dynamics of each reaction, but we only try to capture the global pattern of reaction ratios of an observed dynamics. In other words, leaving unknown the *real* local internal dynamics, we decide to consider the system at an abstraction level which is sufficient to reveal the logic of the behavior we observe. This more abstract approach can be less informative, with respect to specific important details, but such a more generic information could be very useful in discriminating important aspects of the reality, and often, especially in the case of very complex systems, is the only way for grasping a kind of comprehension of the reality under investigation.

We call the system (2) ADA (Avogadro and Dalton Action), when we search to determine $U[i]$ from the knowledge of substance quantities (Avogadro refers to the integer stoichiometric coefficients, and Dalton to the summation of the effects of reactions). The log-gain principle assists us by adding new knowledge to the stoichiometric information of ADA equations. This principle derives from a general biological principle called *allometry* [1], according to which, in a living organism, the global variation of its typical variables are proportional to the relative variations of the variables related to them. In differential terms the relative variation in time of a variable coincides with the variation of its logarithm, therefore we used the term "log-gain" for any law grounded on this assumption. In the specific context of our problem, we assume that *the relative variation of a reaction flux is a linear combination of the relative variations of substance quantities and parameters affecting the reaction.* We refer to the papers [8, 10] for a detailed account on the log-gain theory of MP systems.

The *Log-Gain Stoichiometric Stepwise regression* algorithm (LGSS), presented and motivated in [13], combines and extends the log-gain principle with the classical method of *Stepwise Regression* [6, 2], which is a statistical regression technique based on *Least Square Approximation* [7, 12] and a *Fisher test F*. In fact stepwise regression tries to find the best combination of some prefixed basic functions for approximating a given time series. In LGSS we add the specific knowledge of the stoichiometry of the system under investigation and the requirement that the log-gain principle has to be satisfied in the best possible way. We do not give here the details of the algorithm, which was implemented by suitable MATLAB functions, but it turned definitively out that the addition of these two aspects, related to the particular nature of metabolism, provided an effective improvement of the approximation performance of stepwise regression.

## 3 Statistical distribution of mitotic MP models

In general, LGSS derives MP models from the time series of observed dynamics. In the case of Golbeter's oscillator [3, 4, 5] we found that by considering different values of the resolution time $\tau$, different analytical forms of regulation maps were appropriate. By means of a suitable MATLAB implementation of LGSS, we automatically generated

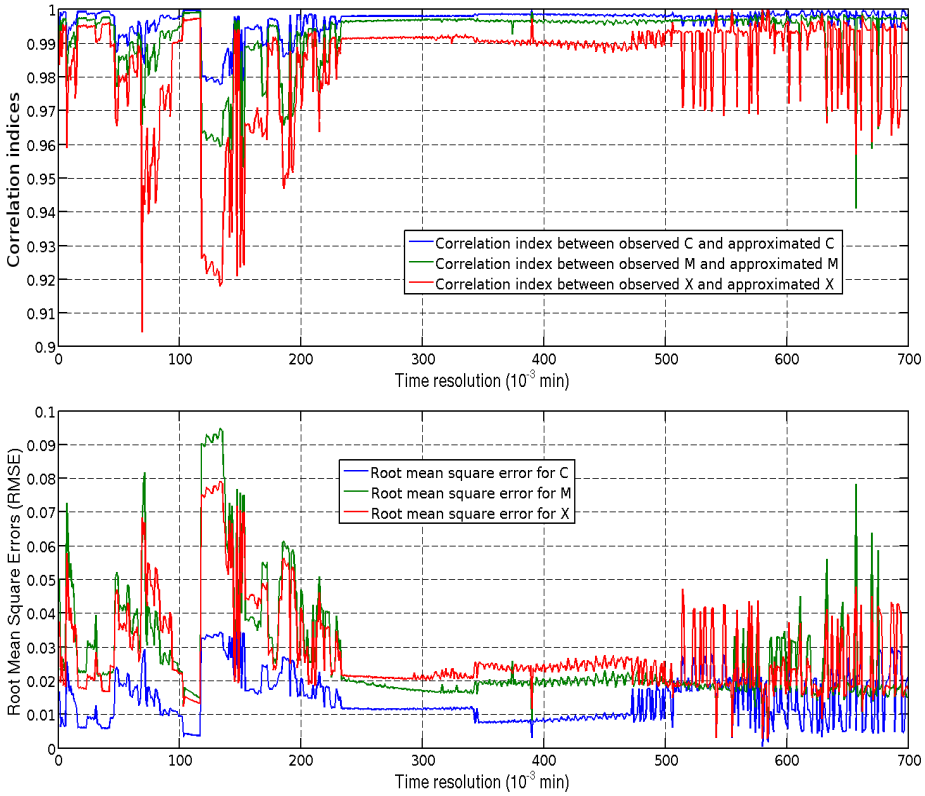



**Fig. 3.** Correlation indices of the three substances for each model w. r. t. the values of resolution time $\tau$ (on the top). Root mean square errors (RMSE) of the three substances for each model w. r. t. the values of resolution time $\tau$ (on the bottom).

700 MP models ($\tau$ varying from $10^{-3}$ min to $700 \cdot 10^{-3}$ min with increments of $10^{-3}$ min). Figure 3 displays Pearson's correlation indices [16] and root mean square errors (RMSE) for each of these models. Each model provides an RMSE with magnitude order at most equal to $10^{-2}$ and permits the calculation of values which are highly correlated to the observed one.

The regulation maps calculated by the LGSS are obtained starting from a dictionary of 20 possible *regressors*, that is monomials of $C$, $M$ and $X$ with degree less than or equal to 3 (i.e. the constant, $C$, $M$, $X$, $C^2$, $M^2$, $X^2$, $CM$, $CX$, $MX$, $C^3$, $M^3$, $X^3$, $C^2M$, $CM^2$, $C^2X$, $CX^2$, $M^2X$, $MX^2$ and $CMX$)[2]. Figure 4 displays two diagrams giving the number of regressors and the total number of monomials occurring in all the regulation maps of each MP model. We need at least 6 different regressors to get a good MP model while we need at least 18 monomials to define all the regulation maps that comprise a model.

[2] Substances $M^+$ and $X^+$ are not considered because they depend on $M$ and $X$ respectively.

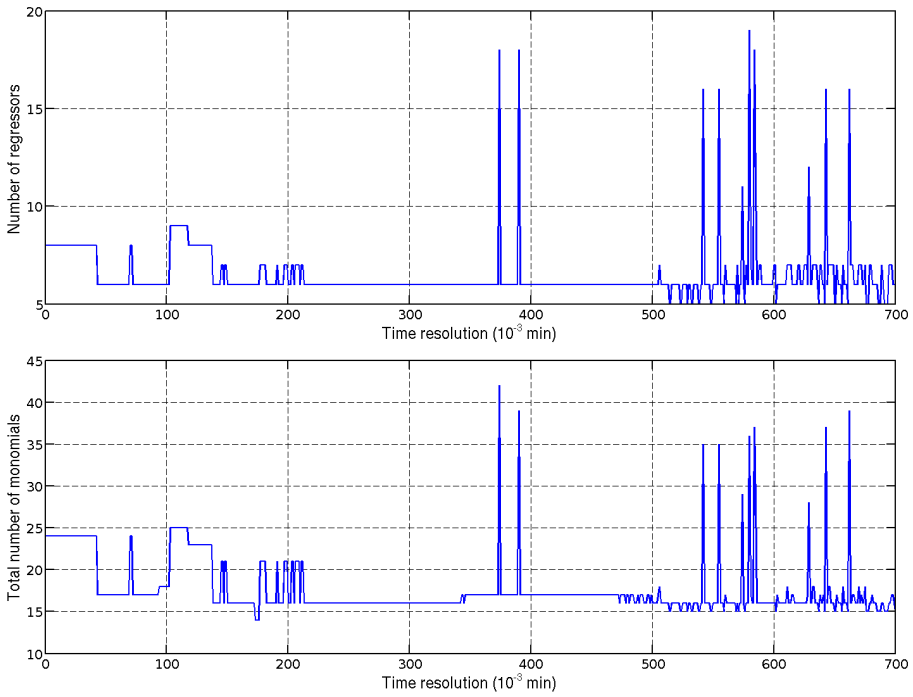



**Fig. 4.** Number of different regressors for each model w. r. t. the values of resolution time $\tau$ (on the top). Total number of monomials for each model w. r. t. the values of resolution time $\tau$ (on the bottom).

## 4 Model classification according to descriptional parameters

Given a MP grammar, providing the dynamics of a MP model, we can abstract from the particular values of the constants of regressors by identifying a MP grammatical schemata defined in terms only of the analytical form of regressors constituting each regulation map. In this section we present the distribution of these grammatical schemata over the population of mitotic models. We found that all the 700 models are distributed into 40 different grammatical schemata. If we order these schemata according to the number of models where they occur we find the distribution given in Figure 5. For example the grammatical schemata at the 10th position has 26 models. In Table 2 other descriptional indices of models are given for the first 14 grammatical schemata which define 621 models from a total of 700 (89%). These indices are useful for discriminating interesting aspects of the MP grammars and they comprehend:

1. the *number of regressors*;
2. the *total number of monomials*;
3. the *temporal grain* of dynamics observation which is expressed by the values of time interval $\tau$;

4. the *best value of* $\tau$ which is relative to the model which provides the best dynamical approximation of the mitotic phenomenon;
5. the best *RMSE* which is the average value of the RMSE relative to the substances curves corresponding to the best $\tau$.

It is worthwhile to remark that the grammatical schemata occurring at the first positions (with high frequency) are also the grammatical schemata having a small number of regressors and monomials.

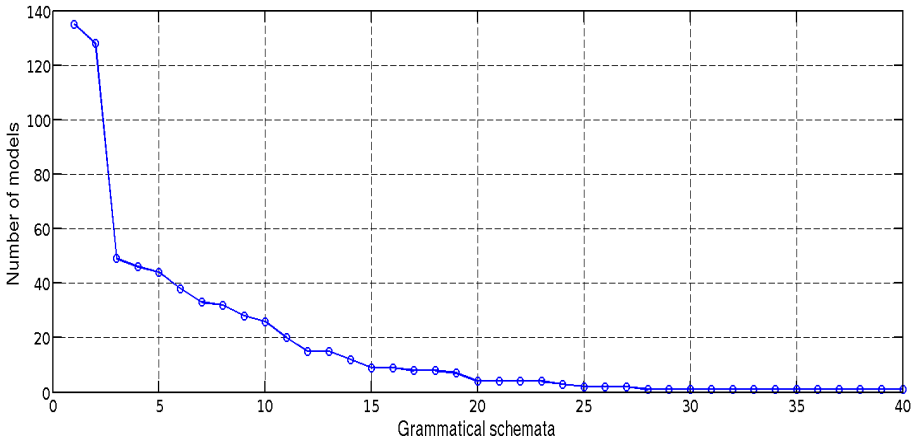



**Fig. 5.** Number of MP models w. r. t. grammatical schemata (please see Table 2 for details).

| Grammatical schemata | number of models | number of regressors | total n. of monomials | $\tau$ interval $(10^{-3}\ min)$ | best $\tau$ $(10^{-3}\ min)$ | best RMSE |
|---|---|---|---|---|---|---|
| 1 | 135 | 6 | 16 | 151 – 345 | 315 | $1.61 \cdot 10^{-2}$ |
| 2 | 128 | 6 | 17 | 343 – 477 | 401 | $1.62 \cdot 10^{-2}$ |
| 3 | 49 | 6 | 17 | 43 – 93 | 43 | $1.84 \cdot 10^{-2}$ |
| 4 | 46 | 6 | 16 | 138 – 232 | 219 | $1.95 \cdot 10^{-2}$ |
| 5 | 44 | 8 | 24 | 1 – 71 | 40 | $1.48 \cdot 10^{-2}$ |
| 6 | 38 | 6 | 16 | 525 – 699 | 683 | $1.78 \cdot 10^{-2}$ |
| 7 | 33 | 6 | 16 | 473 – 563 | 556 | $1.79 \cdot 10^{-2}$ |
| 8 | 32 | 5 | 15 | 514 – 694 | 602 | $2.78 \cdot 10^{-2}$ |
| 9 | 28 | 7 | 16 | 570 – 696 | 671 | $1.09 \cdot 10^{-2}$ |
| 10 | 26 | 6 | 16 | 493 – 684 | 684 | $1.8 \cdot 10^{-2}$ |
| 11 | 20 | 8 | 23 | 118 – 137 | 137 | $5.86 \cdot 10^{-2}$ |
| 12 | 15 | 9 | 25 | 103 – 117 | 103 | $9.6 \cdot 10^{-3}$ |
| 13 | 15 | 6 | 17 | 474 – 499 | 474 | $1.62 \cdot 10^{-2}$ |
| 14 | 12 | 7 | 21 | 191 – 212 | 212 | $1.97 \cdot 10^{-2}$ |

**Table 2.** Descriptional indices of models given for the first 14 grammatical schemata ordered as explained in section 4.

### 4.1 Analytical forms of mitotic MP grammars

In this section we will present a number of MP models where regulation maps present very simple and nice forms especially in comparison with the analytical form of Goldbeter's model of Table 1. They were chosen according to some criteria of representativeness that are based on the classification analysis developed in the previous section.

Tables 3 and 4 give the best two MP mitotic oscillators whose grammars belong to the first two grammatical schemata presented in Table 2. Tables 5 and 6 give the two simpler MP grammars which define a mitotic oscillator: the first one uses the minimum number of different regressors (only 5: the constant, $C$, $M$, $X$ and $CM$) while the second one uses the minimum (total) number of monomials (only 14 monomials).

$$
\begin{array}{ll}
r_1 : \emptyset \rightarrow C & \varphi_1 = v_i \\
r_2 : C \rightarrow \emptyset & \varphi_2 = k_1 + k_2\ C + k_3\ M + k_4\ X - k_5\ C^2 - k_6\ CM \\
r_3 : M^+ \rightarrow M & \varphi_3 = k_7 + k_8\ CM \\
r_4 : M \rightarrow M^+ & \varphi_4 = k_9\ M + k_{10}\ X \\
r_5 : X^+ \rightarrow X & \varphi_5 = k_{11}\ C + k_{12}\ M \\
r_6 : X \rightarrow X^+ & \varphi_6 = k_{13} + k_{14}\ X + k_{15}\ C^2 + k_{16}\ CM
\end{array}
$$

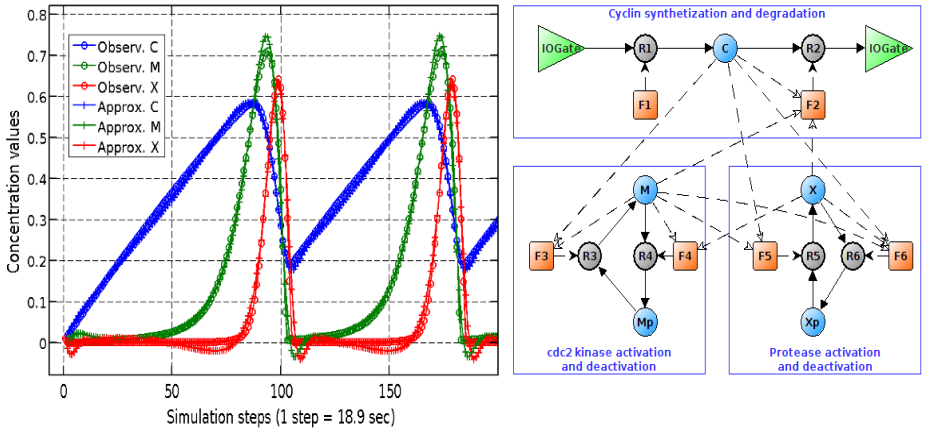



**Table 3.** The best MP mitotic oscillator whose MP grammar belongs to the grammatical schemata 1 of Table 2 ($\tau = 315 \cdot 10^{-3}$ min, $RMSE \approx 1.61 \cdot 10^{-2}$). Constants and initial values: $v_i = 0.025$, $k_1 = 0.0158$, $k_2 = 0.0168923$, $k_3 = 0.0428226$, $k_4 = 0.054506$, $k_5 = 0.03327$, $k_6 = 0.0485192$, $k_7 = 0.00245843$, $k_8 = 0.540636$, $k_9 = 0.219284$, $k_{10} = 0.14129$, $k_{11} = 0.308615$, $k_{12} = 1.01307$, $k_{13} = 0.0338141$, $k_{14} = 0.468994$, $k_{15} = 0.756053$, $k_{16} = 1.15991$, $C[0] = M[0] = X[0] = 0.01$, $M^+[0] = X^+[0] = 0.99$.

$$
\begin{array}{ll}
r_1 : \emptyset \rightarrow C & \varphi_1 = v_i \\
r_2 : C \rightarrow \emptyset & \varphi_2 = k_1 + k_2\ C + k_3\ M + k_4\ X - k_5\ C^2 - k_6\ CM \\
r_3 : M^+ \rightarrow M & \varphi_3 = k_7\ C + k_8\ CM \\
r_4 : M \rightarrow M^+ & \varphi_4 = k_9 + k_{10}\ M + k_{11}\ X \\
r_5 : X^+ \rightarrow X & \varphi_5 = k_{12}\ C + k_{13}\ M \\
r_6 : X \rightarrow X^+ & \varphi_6 = k_{14} + k_{15}\ X + k_{16}\ C^2 + k_{17}\ CM
\end{array}
$$

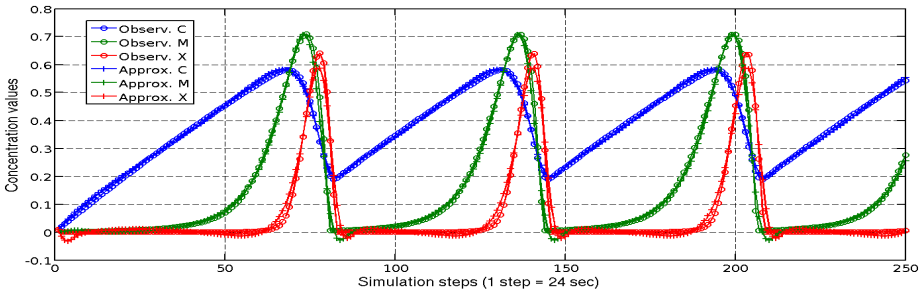



**Table 4.** The best MP mitotic oscillator whose MP grammar belongs to the grammatical schemata 2 of Table 2 ($\tau = 401{\cdot}10^{-3}$ min, $RMSE \approx 1.62{\cdot}10^{-2}$). Constants and initial values: $v_i = 0.025$, $k_1 = 0.0129$, $k_2 = 0.0255671$, $k_3 = 0.0666719$, $k_4 = 0.0632731$, $k_5 = 0.0522867$, $k_6 = 0.0749538$, $k_7 = 0.02125$, $k_8 = 0.836282$, $k_9 = 0.00202831$, $k_{10} = 0.385222$, $k_{11} = 0.14451$, $k_{12} = 0.392585$, $k_{13} = 1.2218$, $k_{14} = 0.0346891$, $k_{15} = 0.586917$, $k_{16} = 0.962714$, $k_{17} = 1.35871$, $C[0] = M[0] = X[0] = 0.01$, $M^+[0] = X^+[0] = 0.99$.

$$
\begin{array}{ll}
r_1 : \emptyset \rightarrow C & \varphi_1 = v_i \\
r_2 : C \rightarrow \emptyset & \varphi_2 = k_1 + k_2\ M + k_3\ X - k_4\ C - k_5\ CM \\
r_3 : M^+ \rightarrow M & \varphi_3 = k_6\ C + k_7\ CM \\
r_4 : M \rightarrow M^+ & \varphi_4 = k_8 + k_9\ M + k_{10}\ X \\
r_5 : X^+ \rightarrow X & \varphi_5 = k_{11} + k_{12}\ M \\
r_6 : X \rightarrow X^+ & \varphi_6 = k_{13}\ C + k_{14}\ X + k_{15}\ CM
\end{array}
$$

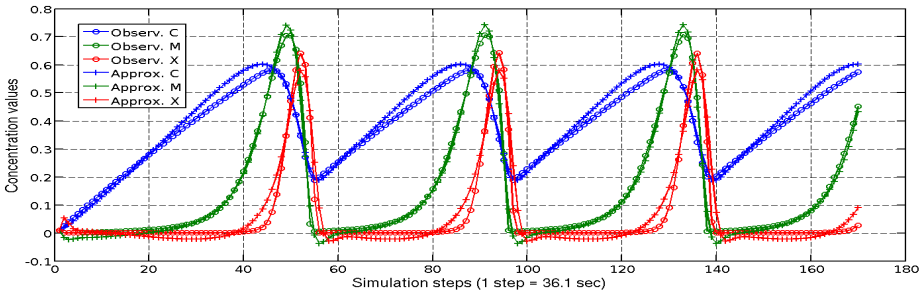



**Table 5.** The MP mitotic oscillator with the minimum number of different regressors ($\tau = 602 \cdot 10^{-3}$ min, $RMSE \approx 2.78 \cdot 10^{-2}$). Constants and initial values: $v_i = 0.025$, $k_1 = 0.0123$, $k_2 = 0.116301$, $k_3 = 0.0922507$, $k_4 = 0.00704311$, $k_5 = 0.148285$, $k_6 = 0.0596357$, $k_7 = 1.78159$, $k_8 = 0.0162002$, $k_9 = 0.922378$, $k_{10} = 0.119154$, $k_{11} = 0.0388314$, $k_{12} = 1.38018$, $k_{13} = 0.173718$, $k_{14} = 0.634806$, $k_{15} = 1.69501$, $C[0] = M[0] = X[0] = 0.01$, $M^+[0] = X^+[0] = 0.99$.

$$
\begin{array}{ll}
r_1 : \emptyset \rightarrow C & \varphi_1 = v_i \\
r_2 : C \rightarrow \emptyset & \varphi_2 = k_1 + k_2\ M + k_3\ X - k_4\ CM \\
r_3 : M^+ \rightarrow M & \varphi_3 = k_5 + k_6\ CM \\
r_4 : M \rightarrow M^+ & \varphi_4 = k_7\ M + k_8\ X \\
r_5 : X^+ \rightarrow X & \varphi_5 = k_9\ C + k_{10}\ M \\
r_6 : X \rightarrow X^+ & \varphi_6 = k_{11} + k_{12}\ X + k_{13}\ C^2 + k_{14}\ CM
\end{array}
$$

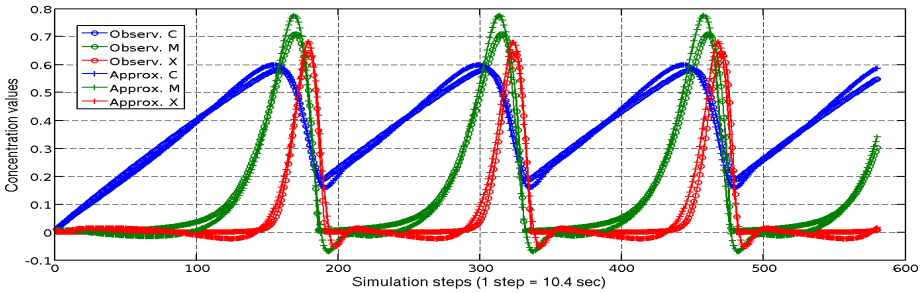



**Table 6.** The MP mitotic oscillator with the minimum total number of monomials ($\tau = 173 \cdot 10^{-3}$ min, *RMSE* $\approx 2.67 \cdot 10^{-2}$). Constants and initial values: $v_i = 0.025$, $k_1 = 0.0209$, $k_2 = 0.0149329$, $k_3 = 0.0351323$, $k_4 = 0.0200062$, $k_5 = 0.000662743$, $k_6 = 0.215816$, $k_7 = 0.0696881$, $k_8 = 0.0911799$, $k_9 = 0.166106$, $k_{10} = 0.569463$, $k_{11} = 0.00823672$, $k_{12} = 0.252676$, $k_{13} = 0.404647$, $k_{14} = 0.668527$, $C[0] = M[0] = X[0] = 0.01$, $M^+[0] = X^+[0] = 0.99$.

## 5 Conclusions

In this paper, by using Golbeter's oscillator as a case study, we show that metabolic P systems yield a robust method for biological modeling. The method we used can be applied without any knowledge about reaction rate kinetics, and can provide, with respect to differential models, different and even simpler mathematical formulations. This possibility is strictly related to the chosen time scale of observed dynamics, and seems to be a promising perspective towards *multi-scale modeling*, which is a challenging aspect in systems biology.

In [13] we develop a systematic analysis and a generalization of the algorithm of Log-Gain Stoichiometric Stepwise Regression (LGSS) on which our results are based. It combines the equational formulation of MP dynamics with the log-gain principle and with the classical statistical regression technique of stepwise regression. This algorithm represents the most recent solution, in terms of MP systems, of the inverse dynamics problem, that is, of the identification of (discrete) mathematical models exhibiting an observed dynamics and satisfying all the constraints imposed by the specific knowledge about the modeled phenomenon.

# Bibliography


[1] Bertalanffy, L.von: General Systems Theory: Foundations, Developments, Applications. George Braziller Inc., New York (1967).

[2] Draper, N., Smith, H.: Applied Regression Analysis, 2nd Edition. John Wiley & Sons, New York (1981).

[3] Goldbeter, A.: A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. PNAS 88(20), 9107–9111 (1991).

[4] Goldbeter, A.: Biochemical Oscillations and Cellular Rhythms: The molecular bases of periodic and chaotic behaviour. Cambridge University Press, Cambridge (1996).

[5] Goldbeter, A.: Computational approaches to cellular rhythms. Nature 420, 238–245 (2002).

[6] Hocking, R.R.: The Analysis and Selection of Variables in Linear Regression. Biometrics 32 (1976).

[7] Luenberger, D.G.: Optimization by Vector Analysis Methods. John Wiley & Sons (1969).

[8] Manca, V.: Log-Gain Principles for Metabolic P Systems. In Condon, A. et al. (eds), Algorithmic Bioprocesses, Natural Computing Series, chapter 28, pp. 585–605, Springer-Verlag (2009).

[9] Manca, V.: Fundamentals of Metabolic P Systems. In [15], chapter 19, Oxford University Press (2010).

[10] Manca, V.: Metabolic P systems. Scholarpedia 5(3):9273 (2010).

[11] Manca, V., Bianco, L., Fontana, F.: Evolutions and Oscillations of P systems: Theoretical Considerations and Application to biological phenomena. In Membrane Computing, WMC 2004, LNCS 3365, pp. 63–84, Springer (2005).

[12] Manca, V., Marchetti, L.: Metabolic approximation of real periodical functions. Journal of Logic and Algebraic Programming, doi:10.1016/j.jlap.2010.03.005 (2010).

[13] Manca, V., Marchetti, L.: Log-Gain Stoichiometic Stepwise regression for MP systems. IJFCS, to appear (2010).

[14] Păun, G.: Membrane Computing. An Introduction. Springer (2002).

[15] Păun, G., Rozenberg, G., Salomaa, A. (eds): Oxford Handbook of Membrane Computing. Oxford University Press (2010).

[16] Pearson, K.: Notes on the History of Correlation. Biometrika 13(1), 25–45 (1920).