Cellular Automata Approaches to Enzymatic Reaction Networks

Jörg R. Weimar

Institute of Scientific Computing, Technical University Braunschweig,
D-38092 Braunschweig, Germany
J.Weimar@tu-bs.de, http://www.jweimar.de

Abstract. Cellular automata simulations for enzymatic reaction networks differ from other models for reaction-diffusion systems, since enzymes and metabolites have very different properties. This paper presents a model where each lattice site can can contain at most one enzyme molecule, but many metabolite molecules. The rules are constructed to conform to the Michaelis-Menten kinetics by modeling the underlying mechanism of enzymatic conversion. Different possible approaches to rule construction are presented and analyzed, and simulations are shown for single reactions and simple enzyme networks.

1 Enzymatic Reaction Networks

Most reactions in biological systems are catalyzed by enzymes. These enzymes are complex molecules (they are proteins) which are not consumed in the reaction, but simply facilitate the reaction of smaller molecules called metabolites (such as sugar). In a biological cell, thousands of different enzymes are active. Each enzymatic reaction takes molecules of one or more metabolite species, the substrates of this reaction, and converts them into molecules of one or more other species, the products. The products are again substrates to other reactions, and thus the reactions form complex networks. Enzymatic reactions can be described and modeled on different levels of detail:

Static Interaction Networks: A first level is the static interaction network, such as the pathways collected in the KEGG database [9, 10, 11], or in the Boehringer-Mannheim map [13]. Such interaction networks can be constructed from purely qualitative information without relying on any quantitative information, such as reaction rates (except possibly for stoichiometric coefficients). Some analyses can extract additional information from these networks, such as inferring elementary metabolic flux modes [14, 15] (although these rely on the information whether a given reaction is reversible or not, which in turn is a semi-quantitative information on the order of magnitude of the equilibrium constant). Simply on the basis of such interaction networks, no quantitative time-dependent simulation is possible.

Reactive Networks: Once quantitative data is available, reaction rates for reactions such as the conversion of a substrate S to a product P by an enzyme E can be described by some rate law, usually in the form of a Michaelis-Menten (MM) law:

$$\frac{d[S]}{dt} = \frac{V_{\max}[E][S]}{K_m + [S]} \tag{1}$$

with the maximum conversion rate $V_{\max}$ and the Michaelis-Menten coefficient K_m.

A more detailed description would contain rates for the elementary reactions leading to such an overall MM rate [1]:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightarrow} E + P. \tag{2}$$

Here, several rate constants need to be measured, which is usually not done or not possible. Most reactions in a realistic reaction network are more complicated, since they involve two or more substrates and two or more products, such as energy-providing ATP. In these cases, more coefficients need to be specified, and in addition the mechanism must be specified, such as BiBiRandom (in which case the two substrates may bind in any order) or Ordered BiBi (where the substrates must bind in a specific order). Other effects, such as competitive inhibition complicate the situation further.

Given a network of enzymatic reactions and the corresponding reaction rates, one can simulate the network by solving the ODEs numerically, or using stochastic simulation methods [7, 8]. A static approach is to analyze steady states, parameter dependences or sensitivities, etc..

Space and Transport Phenomena. The space and diffusion or other transport phenomena are usually only taken into account by compartmentalizing the system where necessary, but they can also be included explicitly (as reaction-diffusion equations, or probabilistically in the stochastic simulation approach [6, 17]). The most detailed simulation would be a full molecular dynamics simulation of the cell, but this is by far not feasible yet.

In this paper I present an approach based on cellular automata for simulating enzymatic reaction networks including diffusive transport.

2 Scales

To establish the conditions for a simulation model, let us first look at the scales of space, time, diffusion coefficients, and molecule concentrations and numbers. First, concentrations and counts: The volume of a typical procaryotic cell is $10^{-15}l$, the typical diameter is $1\mu m = 10000$Å. The diameter of a typical enzyme is 100Å, while the diameter of a typical metabolite is 6Å and of one atom 2Å. The concentrations of the involved molecular species differ widely. For metabolites they can be between $1\mu M$ and $5mM$, while for enzymes they can be in the

range of $10nM \ldots 1\mu M$. This means that a complete cell with volume $10^{-18}m^3$, we find between 1 and 60000 molecules of a given enzyme species, and between $10^5 and 10^9$ molecules of a given metabolite species. Simply from the size of the enzyme, we can calculate that in a space volume of size $(100\text{Å})^3$, there can be O(1) enzyme molecules, but many metabolite molecules. In this paper I consider a simulation with lattice sites which have approximately this size and can contain at most one enzyme molecule of any species (exclusion principle) and an arbitrary number of metabolite molecules. Since we use integers to count the molecules instead of calculating with concentrations, our results will be more similar to the stochastic models than to the PDE models.

The diffusion coefficient for a typical metabolite is around $10^{-9}\frac{m^2}{s}$, for an enzyme that is not bound to a membrane it varies, but is at least two orders of magnitude smaller. The time scales of events inside a cell vary greatly, from $10^{-12}s$ (dissociation events) to $100s$ (fastest cell division). In the cellular automaton model, our diffusion method dictates a connection between time and space scales and diffusion coefficients. In our case, a lattice spacing of 100 Å and typical metabolite diffusion coefficients imply a time step of around $2.5 * 10^{-8}s$, which means that millions of time steps are necessary to simulate dynamical changes of metabolite concentrations.

3 Modeling Enzymes

In our model, each lattice site (we do not use the term cell as usually used in cellular automata to avoid confusion with the biological cell) can contain at most one enzyme molecule but many metabolites. We first consider the options for modeling the enzymatic reaction in one such site.

We show in section 3.1 that directly using the Michealis-Menten rate law is not possible in the discretized setting of cellular automata, then demonstrate in section 3.2 how to obtain correct results by directly simulating the mechanism that was approximated by the Michealis-Menten rate law.

3.1 Michaelis-Menten Rate Law

The first possibility is to directly use a Michaelis-Menten rate law. The rate law in Eq. (1) contains the concentration of the enzyme $[E]$, which in one site is always zero or one. If we assume that the concentrations of metabolites can be simply calculated from the metabolite numbers present in the cell, we have the following cellular automata rule: In time Δt, the probability of converting one molecule S to a molecule P is zero if there is no enzyme E present, otherwise

$$Q_{S\to P}(S) = \frac{\Delta t}{\alpha}\, V_{\max}\frac{\alpha S}{K_m + \alpha S}, \tag{3}$$

where α is a scaling factor to convert molecule counts per lattice site into concentrations. Here we will show that in our setting this method does not give correct results. The formula Eq. 3 is only correct in the limit $\Delta t \to 0$, so that at

most one reaction takes place in one time step. On the other hand, in this limit the Michaelis-Menten simplification is not valid, but should be replaced by the more detailed description in Eq. (2). Another problem appears if the counts S are small (e.g., because concentrations $[S]$ are low or the volume of one lattice size is small). In the case where S is rarely greater than one, which happens when $K_m \ll \alpha$, Eq. (3) reduces to three cases for the probability $Q_{S\to P}$:

$$\begin{aligned} Q_{S\to P}(0) &= 0; \\ Q_{S\to P}(1) &= \frac{\Delta t}{\alpha}\ V_{\max}\frac{\alpha}{K_m+\alpha} = \beta; \\ Q_{S\to P}(S>1) &\approx \frac{\Delta t}{\alpha}\ V_{\max}. \end{aligned}$$

This one-parameter equation described only by β is clearly not a good representation of the original MM law. In fact, we can calculate the average rate of conversion $S \to P$ as a function of the concentration $[S]$ if we assume that the numbers of molecules S at a lattice site follow a Poisson distribution:

$$P(S=k) = \frac{\lambda^k}{k!}e^{-\lambda} \tag{4}$$

with $\lambda = [S]/\alpha$. This distribution has the mean λ and is attained in the absence of reactions and if no limits are imposed on the number of metabolite molecules per cell (if there were limits, it would be a binomial distribution).

On average, the number of molecules transformed is the product of the probability of finding k molecules with the probability of converting one of them, summed over all values k:

$$\langle Q\rangle = \sum_{k=0}^{\infty} P(S=k)\ \ Q_{S\to P}(k) \tag{5a}$$

$$= V_{\max}\ e^{-\lambda}\ (-\lambda)^{-K_m/\alpha}\left(\Gamma\left(\frac{\alpha+K_m}{\alpha},-\lambda\right)-\Gamma\left(\frac{\alpha+K_m}{\alpha}\right)\right), \tag{5b}$$

which for $K_m \gg 1$ is almost equal to the correct MM rate, whereas for small K_m, it is too small (here Γ is the Gamma-function). In Figure 1 we can see a comparison of the exact rate with the average of 100 trials and the predicted average from Eq. (5).

An attempt to correct the discrepancy by changing the function $Q_{S\to P}$ does not lead to an exact solution, since we have some restrictions: Q must be non-negative for all k. To make an exact fit, we would be required to use an alternating diverging series for $Q(k)$. Note that for polynomial reaction rate laws, such a correction is possible and can be given in analytic form [19], while for the MM rate law (which is not polynomial) this is not possible.

3.2 Detailed simulation of the Mechanism

A second possibility is to model the mechanism that leads to the MM rate law in detail. For the simple unimolecular reaction, this mechanism is described by

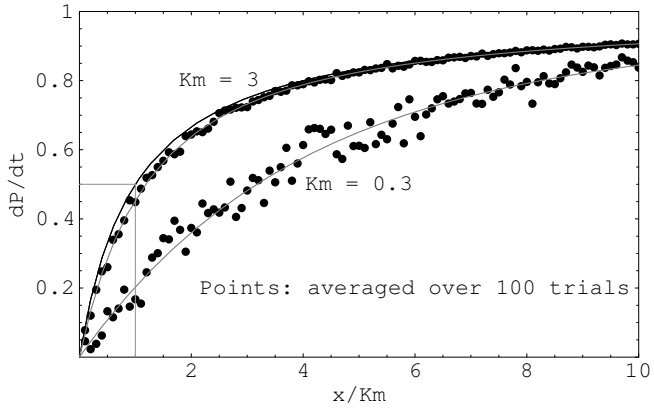

Fig. 1. Correct MM reaction rate (re-scaled by K_m, with $V_{\max} = 1, dt = 1$) (top curve) compared to the average of 100 random realizations (dots) and predicted average from Eq. (5) (two gray curves below the top curve).

Eq. (2). We can directly model this in the automaton, since at each lattice site we have at most one enzyme molecule, and therefore we can keep track of the state of this molecule: unbound or bound to S. The transitions between the states are governed by simple linear rate laws, which we can model exactly in the CA. We therefore have the following processes:

$$\begin{aligned} E \quad + nS &\rightarrow ES + (n-1)S \qquad \text{with probability } p_1(n); \\ ES + nS &\rightarrow E \quad + (n+1)S \qquad \text{with probability } p_{-1}; \\ ES + nP &\rightarrow E \quad + (n+1)P \qquad \text{with probability } p_2, \end{aligned}$$

where the probabilities can be calculated from the rate laws as $p_1(n) = \Delta t\ k_1\ n$, $p_{-1} = \Delta t\ k_{-1}$, $p_2 = \Delta t\ k_2$ and p_{-1} and p_2 are independent of the metabolite numbers present in the cell. One problem with this approach is that the rates k are usually not known. We can assume values by taking into account that $V_m = k_2$ and $K_m = (k_{-1} + k_2)/k_1$ and that k_2 is smaller than k_1 and k_{-1}. If we simply assume $k_{-1} = c\ k_2$, we obtain $k_1 = (c+1)\ V_m/K_m$. One would assume that c should be > 1, but actually $c = 1$ works just as well. Note that now the fastest processes are $(c+1)$ times faster than before, which means that we need to use a smaller time step. With such a model we can perform a number of time steps and show that the average rate of transformation from S to P converges to the rate given by the MM law for all values of K_m.

4 Diffusion

Diffusion of the metabolite molecules can be simulated using a number of techniques. Since we model individual molecule counts, we have to make sure that the number of molecules is conserved by the diffusive operation. This is not

easily verified in standard cellular automata. Two techniques have been developed which make conservation of particles easy in cellular automata: Partitioned cellular automata and block cellular automata.

The first kind is used in the lattice gas automata [2, 4, 5, 18, 19], where each lattice site has a number of channels to the neighboring sites, which can be occupied by at most one particle at a time. Since the transport of particles through these channels in the synchronous cellular automaton simply represents a permutation of the channel contents, one can easily verify that no particles get lost (assuming correct treatment of the boundaries).

The second technique is to subdivide the lattice into blocks, and at each step do some exchange of the contents of the sites within one block [3, 12, 16, 20]. Since only a few cells are involved, it is easy to verify whether conservation laws are observed. In subsequent time steps, the block boundaries are changed to make information exchange across the whole lattice possible. Note that such block cellular automata are equivalent to classical cellular automata, since each can be simulated by the other (with some duplication of cell content and extension of the neighborhood).

Here we use this second technique to simulate diffusion. We use blocks of size two, which are placed on the lattice in all possible orientations in subsequent time steps (four orientations in two dimensions, six in three dimensions). For the exchange between the two cells within a block one can use different prescriptions: One possibility was already mentioned by Gillespie [7, 8] as an extension of his stochastic simulation method to spatially distributed systems: If the two cells contain n_1 and n_2 particles respectively, move $(n_1 - n_2)/2$ particles from cell 1 to cell 2 (or reverse, if the difference is negative). This approach models the macroscopic diffusive flux proportional to the gradient in concentration. Better suited to the stochastic simulation is a microscopic approach: Consider all the particles in both cells as independent, and let each particle move to the neighboring cell with a fixed probability p. Then the number of particles to be moved from cell 1 to cell 2 is obtained by sampling a binomial distribution with parameters p and n_1, while the number of particles moving from cell 2 to cell 1 is given by a binomial distribution with parameters p and n_2. Thus the number of particles exchanged between the two cells is governed by the difference between two binomial distributions. This approach leads to fluctuating particle numbers in all cells even in the absence of reactions, which is appropriate when particles represent molecules. The particle numbers in the cells will be distributed according to a Poisson distribution Eq. (4).

For the calculations of effective reaction rates, as in Eq. (5), the averaged (macroscopic) diffusion operator leads to a distribution that is more compact than the Poisson distribution, e.g. a two-valued distribution around the average concentration. Here we use the microscopic procedure to ensure consistent results.

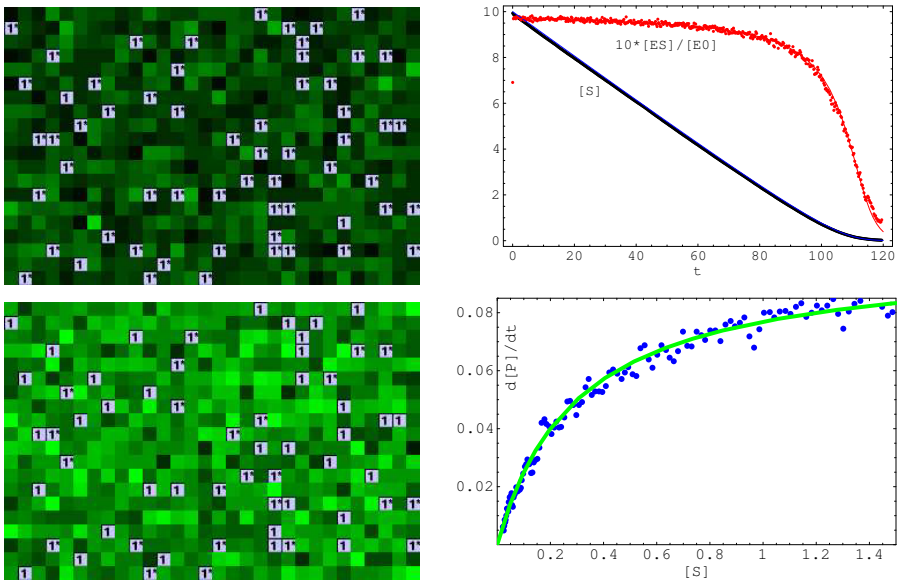

Fig. 2. Simulation of one reaction using the CA. Cells marked by a 1 contain an enzyme molecule, when marked with "1*" the enzyme is bound to a substrate molecule. The other cells show the relative number of P molecules. On the left we show snapshots at time $t = 50\Delta t$ and $t = 110\Delta t$, on the right the time evolution of $[S]$ and $[ES]/[E_0]$ (the fraction of substrate-bound enzyme molecules), and right bottom the dependence of $d[P]/dt$ on $[S]$ (for a larger 3-D system of 30^3 sites). The lines show the theoretical predictions from the MM rate law.

5 CA simulations

Figure 2 shows a simulation with such a cellular automaton for a system with only one enzymatic reaction: $E + S \rightarrow E + P$. We use the parameters $V_m = 1, K_m = 0.3, c = 1, \Delta t = 0.01$. Here the time step is limited by the restriction that at most one molecule can react at one time step (since the enzyme can only bind one substrate molecule at a time), and we start the simulation with a high substrate concentration of 10 molecules per site. We observe that the CA model corresponds to the predictions from the macroscopic Michaelis-Menten rate law. In this example diffusion of the metabolites is comparably fast, and the spatial dimension does not have a measurable influence.

In a second test, we use this model to simulate a toy network of unidirectional enzymatic reactions with four enzymes and four metabolites.

$$A \xrightarrow{E_1} B \xrightarrow{E_2} C \xrightarrow{E_3} D \xrightarrow{E_4} A \tag{6}$$

The parameters for the different enzymes are: E_1: $V_{m1} = 1$, $K_{m1} = 0.3$, E_2: $V_{m2} = 2$, $K_{m2} = 1$, E_3: $V_{m3} = 3$, $K_{m3} = 0.3$, E_4: $V_{m4} = 3$, $K_{m4} = 3$, the initial condition is $[B] = 10$ and $[A] = [C] = [D] = 0$.

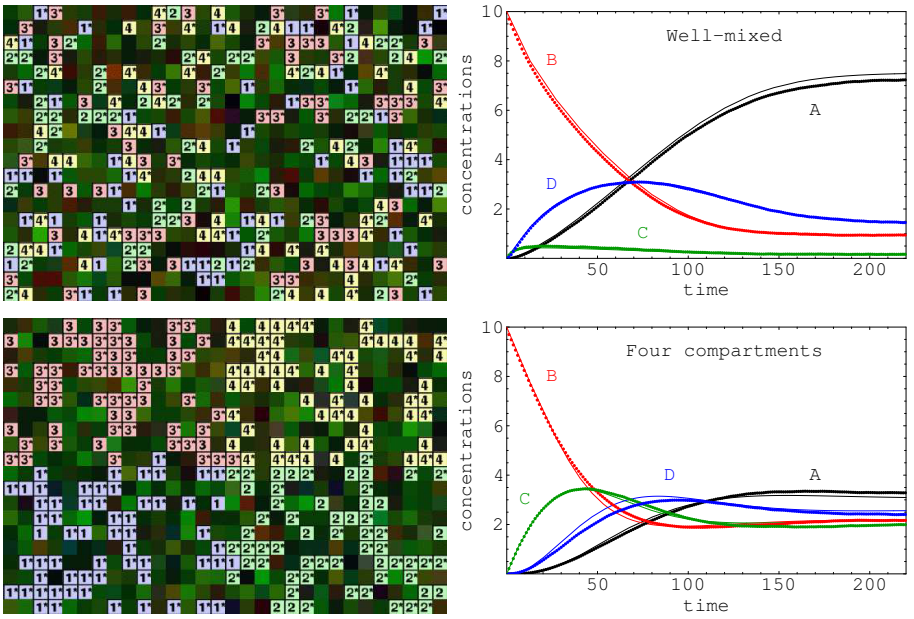

Fig. 3. Simulation of four-enzyme system with different spatial distributions of the enzymes. Top row: random distribution of enzymes (with equal numbers for all enzymes, and probability of 0.2 for each site). Bottom row: distribution in four quadrants. The concentration time series shown at right are compared with the predictions from the differential equations using Michaelis-Menten kinetics. For the four quadrant model, we compare with a compartmentalized ODE model, since the well-mixed ODE model, which is the same in both cases, clearly gives wrong results. The initial conditions for the metabolites are identical in both cases, and the number of enzyme molecules of each type is also identical in both cases.

We simulate two different situations in Figure 3, one where the enzymes are randomly distributed in space, and one where the enzymes are separated in different compartments, and the metabolites must diffuse from one enzyme to the other. In the first case, we compare the time evolution of the metabolite concentrations with the solution of the ordinary differential equations, using the Michaelis-Menten approximation for the construction of the ODEs. In the second case (with identical parameters except for the spatial distribution of enzymes), we observe a significantly different time evolution and also a very different steady state. We can model this situation by constructing a system of ODEs where each quadrant is treated as a different compartment, the reaction rates are adjusted in the compartments (e.g., in the compartment corresponding to the left lower quadrant, only enzyme E_1 is present, but in concentration four times higher than in the first model. Therefore only the reaction $A \rightarrow B$ takes place, with rate four times higher), and diffusion of all species between neighboring compartments is

introduced. The lines in the time evolution plot correspond to the solution of this ODE.

The large differences in the two solutions show that spatial dependences should not be ignored when whole-cell models are constructed.

6 Conclusion

We have described a cellular automaton model for enzymatic reaction networks. This model is based on block-cellular automata to ensure conservation of particles and assumes that at most one enzyme can be present at any lattice site. The enzymes can change state by binding to metabolite molecules, and the state changes are described by probabilistic rules derived from the enzyme kinetics to be simulated. The quantitative correctness of the rules has been demonstrated by analytic arguments and by comparison of simulations with the ODE solution. This model can incorporate complex processes, where an enzyme binds several ligands, possibly in specific order. The explicit inclusion of space makes possible the detailed investigation of phenomena that depend on diffusion of different species. As an example we have demonstrated that simply placing the enzymes at different regions leads to a significantly different average behavior. In this case the behavior can be well approximated by a compartmentalized ODE model, but in more complex geometries of biological relevance, this is not necessarily the case.

The model presented here is a microscopic model, since individual enzyme molecules are explicitly represented, and the corresponding fluctuations take place, which sets the model apart from numerical methods for solving the averaged PDE. We showed that enzymatic reactions cannot be simulated by the same techniques used e.g., in reactive lattice gas automata [2], since the rate law is not polynomial.

As further steps, we will apply this cellular automaton approach to more complex enzymatic reaction networks and try to obtain biologically meaningful spatial distributions of the individual enzymes.

References

[1] Hans Bisswanger. *Enzymkinetik*. Wiley-VCH, Weinheim, 2000.

[2] Jean Pierre Boon, David Dab, Raymond Kapral, and Anna Lawniczak. Lattice gas automata for reactive systems. *Physics Reports*, 273(2):55–147, 1996.

[3] B. Chopard and M. Droz. Cellular automata approach to diffusion problems. In P. Manneville, N. Boccara, G. Y. Vichniac, and R. Bidaux, editors, *Cellular automata and modelling of complex physical systems*, pages 130–143, Berlin, Heidelberg, 1989. Springer-Verlag.

[4] David Dab and Jean-Pierre Boon. Cellular automata approach to reaction-diffusion systems. In P. Manneville, N. Boccara, G. Y. Vichniac, and R. Bidaux, editors, *Cellular automata and modelling of complex physical systems*, pages 257–273, Berlin, Heidelberg, 1989. Springer-Verlag.

[5] David Dab, Jean-Pierre Boon, and Yue-Xian Li. Lattice-gas automata for coupled reaction-diffusion equations. *Phys. Rev. Lett.*, 66(19):2535–2538, 1991.

[6] Micheal A. Gibson. *Computational Methods for Stochastic Biological Systems.* Phd thesis., Calif. Inst. Technology, 2000.

[7] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Computational Physics*, 22:403–434, 1976.

[8] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem*, 81:2340–2361, 1977.

[9] M. Kanehisa. Kegg. http://www.genome.ad.jp/kegg/.

[10] M. Kanehisa. Toward pathway engineering: a new database of genetic and molecular pathways. *Science & Technology Japan*, 59:34–38, 1996.

[11] M. Kanehisa and S. Goto. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, 28:27–30, 2000.

[12] Norman Margolus and Tomaso Toffoli. *Cellular automata machines. A new environment for modeling.* MIT Press, Cambridge, MA, USA, 1987.

[13] Gerhard Michal. *Boehringer Mannheim Biochemical Pathways Map.* Roche and Spektrum Akademischer Verlag.

[14] Stefan Schuster. Studies on the stoichiometric structure of enzymatic reaction systems. *Theory Biosci.*, 118:125–139, 1999.

[15] Stefan Schuster and Claus Hilgetag. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.*, 2:165–182, 1994.

[16] T. Toffoli and N. Margolus. Invertible cellular automata: a review. *Physica D*, 45:229–253, 1990.

[17] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry.* Elsevier, 1992.

[18] Jörg R. Weimar. *Cellular Automata for Reactive Systems.* PhD thesis, Université Libre de Bruxelles, Belgium, 1995.

[19] Jörg R. Weimar. *Simulation with Cellular Automata.* Logos-Verlag, Berlin, 1998.

[20] Jörg R. Weimar. Simulating reaction-diffusion cellular automata with JCASim. In T. Sonar, editor, *Discrete Modelling and Discrete Algorithms in Continuum Mechanics*, pages 217–226. Logos-Verlag, Berlin, 2001.