ORIGINAL PAPER

A stochastic model of autocatalytic reaction networks

**Alessandro Filisetti · Alex Graudenzi · Roberto Serra ·
Marco Villani · Rudolf M. Füchslin · Norman Packard ·
Stuart A. Kauffman · Irene Poli**

Abstract Autocatalytic cycles are rather widespread in
nature and in several theoretical models of catalytic reaction
networks their emergence is hypothesized to be inevitable
when the network is or becomes sufficiently complex.
Nevertheless, the emergence of autocatalytic cycles has
been never observed in wet laboratory experiments. Here,
we present a novel model of catalytic reaction networks
with the explicit goal of filling the gap between theoretical
predictions and experimental findings. The model is based
on previous study of Kauffman, with new features in the
introduction of a stochastic algorithm to describe the
dynamics and in the possibility to increase the number of
elements and reactions according to the dynamical evolu-
tion of the system. Furthermore, the introduction of a
temporal threshold allows the detection of cycles even in
our context of a stochastic model with asynchronous update.
In this study, we describe the model and present results
concerning the effect on the overall dynamics of varying

(a) the average residence time of the elements in the reactor,
(b) both the composition of the firing disk and the con-
centration of the molecules belonging to it, (c) the com-
position of the incoming flux.

Introduction

There are numerous examples of autocatalytic cycles (or
autocatalytic sets, ACS from now on) in the present day
biological systems, the outcome of billions of years of
evolution. In this regard, if one is interested either in the
problem of the origin of life, or in the design of artificial
protocells (Carletti et al. 2008; Filisetti et al. 2008;

A. Filisetti (✉) · A. Graudenzi · R. Serra · M. Villani · I. Poli
European Centre for Living Technology, Calle del Clero 2940,
30124 Venice, Italy
e-mail: alessandro.filisetti@ecltech.org

A. Graudenzi
e-mail: alex.graudenzi@ecltech.org

R. Serra
e-mail: rserra@unimore.it

M. Villani
e-mail: marco.villani@unimore.it

I. Poli
e-mail: irenpoli@unive.it

R. Serra · M. Villani
Department of Communication and Economics, University of
Modena and Reggio Emilia, via Allegri 9,
42100 Reggio Emilia, Italy

R. M. Füchslin
Artificial Intelligence Laboratory, University of Zürich,
Andreasstr. 15, 8050 Zurich, Switzerland
e-mail: fuchslin@ifi.uzh.ch

N. Packard
ProtoLife Inc., 57 Post St. Suite 513, San Francisco, CA 94104,
USA
e-mail: n@protolife.com

S. A. Kauffman
Departments of Biochemistry and Mathematics, University of
Vermont, Burlington, VT 05405, USA
e-mail: Stuart.Kauffman@uvm.edu

I. Poli
Dipartimento di Statistica, Università Ca' Foscari, San Giobbe,
Cannaregio 873, 30121 Venezia, Italy

Rasmussen et al. 2004; Serra et al. 2007; Szostak et al. 2001), the question on the *emergence* of ACSs and the characterization of the generic properties of catalytic reaction networks in general is of fundamental importance. Here, we introduce a novel model tailored for an in-depth investigation of the issue.

Models of this kind have been developed in the past, including the well-known studies by Dyson (1985), Eigen and Schuster (1979), Kauffman (1986), Jain and Khrishna (2001), and Kaneko (2006). These models can be grouped in two classes, depending on the underlying hypotheses about the physical nature of the collectively replicating molecules and of their interactions. The first class is based on the hypothesis of "gene-first" scenarios and is inspired by nucleic acid template matching, where a linear polymer drives the synthesis of its complementary strand; the second class is instead inspired by "protein-first" scenarios, according to which a molecule can catalyze reactions involving other molecules without imposing the constraint of complementary matching.[1]

Note that nowadays there is a wide consensus that the appearance of collectively self-replicating sets of molecules is not per se "life", whose properties require also the existence of a container which separates the "living system" from its environment and a coupling between the dynamics of the replicating molecules and the growth and division of the container. Once such coupling has been achieved, it has been proven that the rates of the two processes (duplication of the genetic molecules and of the container) tend to spontaneously synchronize through successive divisions (Serra et al. 2007; Filisetti et al. 2008; Carletti et al. 2008), leading to an exponential growth of a population of protocells, i.e., a Darwinian selection between them (Munteanu et al. 2006). In any case, the appearance of a set of collectively self-replicating molecules is a crucial step both for the origin of life and for that of an effective protocell.

It is interesting to notice that, in spite of their important differences, all the above mentioned models describe a qualitative change in the system behavior, such that when a certain set of conditions is met at least an ACS emerges. On the other hand, it is observed in wet lab experiments that obtaining a set of molecules endowed with such a property is a very difficult task, and this seems of course to contradict the apparent universality of the theoretical behavior. Our model is aimed to investigate the reasons of such a mismatch between the theoretical predictions and the experimental results.

The first inspiration of this study is one of the above-mentioned models, i.e., proposed by Kauffman (1986),

which is summarized in "Critical review of previous results" section.

In brief, the molecules are linear chains of monomers taken from a finite alphabet. There are two possible reactions, namely condensation (two polymers are joined forming a longer one) and cleavage (a polymer gives rise to two by splitting at a certain point). It is assumed that these reactions occur at a negligible rate unless they are catalyzed, and it is also assumed that any molecule has a small but finite chance to catalyze any random reaction. By enumerating all the possible reactions and molecules, Kauffman came to the conclusion that, provided that there are enough different molecules in the initial set, a connected component will appear in the reaction graph, this corresponding to the presence of (at least) an ACS.

The reasoning of Kauffman about the transition phase is compelling, but the simplifications are drastic. In particular, in his original study Kauffman did not consider selection nor the concentration of the molecules, but focused his attention on the graph of the reactions which are possible among molecules which can be synthesized. Indeed, no extinction, no small-number effects are taken into account. To overcome this setback, a further evolution of the model was introduced by Farmer et al. (1986) and Bagley and Farmer (1991), who switched to a representation in terms of differential equations, thus taking the concentrations into account. Moreover, they modelled a chemostat, thereby providing a kind of selective pressure. Their results are interesting, and will be briefly summarized in "Critical review of previous results" section. However, the use of a continuous formalism does not allow to take stochastic effects into account properly (Bower and Bolouri 2004). In order to partially circumvent this problems, the authors introduced a threshold, roughly corresponding to a molecule per reaction volume,[2] so that when the (continuous valued) concentration falls below that level it is suddenly set to 0 (Bagley and Farmer 1991).

In our model we stick to the original Kauffman representation, but we relax the non-physical constraints regarding a maximum allowed length, allowing the creation of new species and of new reactions according to the dynamics, and we introduce the quantity of the different molecules in the system. The dynamics is described using a stochastic approach, by means of the well-known Gillespie algorithm (Gillespie 1977), to consider the stochastic nature of the dynamics of interacting molecules in real systems, with particular attention to the instability points and the cases in which the number of some of the species is considerably low. Moreover, we avoid three-molecular reactions because of their extreme rarity, splitting the

---

[1] We also mention the lipid-world hypothesis (Segre et al. 1998).

[2] The reaction volume was taken to be similar to that of a small bacterium, i.e. 1 μm^3.

condensation reaction in two steps: first the formation of a complex between the catalyst and a substrate, and later the encounter between the complex and the second substrate. The complex has a finite lifetime, so it may dissociate before condensation occurs.

In previous study (Filisetti et al. 2011), we provided an in-depth formal description of the features of the model and we illustrated some preliminary analysis concerning the composition of the incoming flux in critical systems. In particular, we showed that the number of different species that are present in the flux seriously influences the overall dynamics, with particular regard to the emergence of ACSs (which become more probable as the number of species increases); on the opposite, (limited) variations of the concentration of the singular species appear to play a minor role in the enhancement of the overall activity of the system. Furthermore, we proved that an increase of the average length of the polymers in the incoming flux does not alter the dynamics in a significant way.

The article is organized as follows: In "Critical review of previous results" section a critical review of the results of the previous models of catalytic reaction networks is presented. Our model is then described in detail in "The model" section, where we also deal with the problem of the choice of the appropriate reaction graph, with particular attention to the introduction of a specific temporal threshold that allows to identify cycles also within an asynchronous stochastic update framework. In "Results of the analyses" section we report the results of the analyses regarding some of the key parameters of the system, to understand whether and how their variation may affect the chance of emergence and maintenance of ACSs. In particular, we study the influence of the average residence time of the elements present in the reactor and we compare the effects of variations of the composition of the firing disk, i.e., the initial set of all the species up to a certain length (some species may have an initial concentration equal to 0), and of the incoming flux on the overall dynamics. Notice that in several cases one finds that there are some species with low numerosity whose presence turns out to be crucial. This observation proves the usefulness of choosing a stochastic framework dealing with this issue. In the final section indications for further study are provided.

Critical review of previous results

The main idea contained in the original study proposed in Kauffman (1986) is that the emergence of ACSs of reactions is unavoidable when starting from a mixture containing *enough* polypeptides. Considering polypeptides composed of two monomers A and B and an initial

population in which all polypeptides up to length M are present, the total number of species is

$$S^M = \sum_{L=1}^{M} 2^L = 2^{M+1} - 2. \tag{1}$$

As anticipated in "Introduction" section, we consider two possible reactions: condensation and cleavage, therefore, the total number of reactions building a specific polymer of length $L^*, 1 \leqslant L^* \leqslant M$ is

$$R_{L^*,i}^M = \sum_{i=L+1}^{M} (2 \times 2^{i-L^*}) + (L^* - 1). \tag{2}$$

It is possible to show (Kauffman 1986) that the ratio between the total number of reactions among polymers and the total number of species is equal to

$$\frac{R_{\text{tot}}^M}{S^M} = \sum_{i=1}^{M} \frac{M-i}{2^i} \cong M - 2. \tag{3}$$

Equation 3 shows that, although the total number of polymers increases exponentially, the number of conceivable reactions, (i.e., the number of potential reactions by which polymers up length M can be interconverted) increases even faster, leading to the linear increase of their ratio. Therefore, an ACS will certainly form, no matter what the probability of catalysis is, provided that there are polymers long enough, which in this case corresponds to having a sufficient number of different types of polymers in the system (Kauffman 1986).

Equation 2 shows the number of reactions able to build a specific polymer of length $L^*, 1 \leqslant L^* \leqslant M$. This number decreases as L approaches M, and therefore there are more ways to form short polymers than long polymers. Nevertheless, one should also consider that for any given length L there are 2^L different polymers formed by a 2 letters alphabet; as a consequence the number of reactions which give rise to polymers of length L is

$$R_{L,\text{tot}}^M = 2^L \left[\sum_{i=L+1}^{M} (2 \times 2^{i-L}) + (L-1) \right] - 2^{L+1}. \tag{4}$$

The last term takes into account the case, where the index of the sum is $i = 2L$.[3] Therefore, although there are more ways to create a short specific polymer than a longer one, the formation of long polymers is favored.

The concentrations of the various species are taken into account in the models (Farmer et al. 1986; Bagley and Farmer 1991). In both cases simulations were performed using a deterministic model, i.e., a set of ordinary differential equations.

---

[3] In this case one single reaction is enough to produce two different polymers of equal length, so the total amount of reactions have to be reduced by $2 \times 2^{2L-L} = 2^{L+1}$ reactions.

Analyses were based on the connectivity ratio between edges and nodes, R, of the reactions graph. If $R < 0.5$ the system is *subcritical* and no ACS is formed, between 0.5 and 1 a critical zone is established and some structures begin to form and $R > 1$ corresponds to the *supracritical* state, where cycles appear. In particular, they investigated the relations between the probability that a randomly chosen molecule catalyzes a randomly chosen reaction, and the properties of the initial set of species, pointing out the importance of providing a rich set of different species to allow the formation of cycles. Note that in this model, as in Kauffman's, the initial food set is composed of all the species up to length M, so it is impossible to distinguish the effects of the numerosity from that of the maximum allowed length. If we call B the size of the alphabet and L the maximum length of the firing disk, criticality is achieved if $p \approx B^{-2L}$. A well stirred chemostat having the same dimension of a bacterium is simulated by Bagley and Farmer ([1991](#)).

The model

Main features

For an exhaustive description of the model please refer to Filisetti et al. ([2011](#)). Here, we will outline its main features. The aim of the model is not to provide a detailed description of a specific set of reactions; rather, it is to focus on the general dynamics emerging from the interaction of a large number of interacting molecules.

The basic entities of this model are linear chains of letters from a finite alphabet. The model is fairly general and does not refer to specific chemical classes; in particular, the basic units could represent single elements, stable compounds, or more complex domains. Given their essential characteristic of allowing recursive assembly, let us call them (without losing generality) "monomers", whereas the linear chains could be called "polymers". Species will be used to denote either monomers or polymers.

Species can change in time and they will be denoted at a given time by $(x_1, x_2, \ldots, x_n)$. In the sequence of monomers the order matters, reading from left to right (i.e., *ABB* is different from *BBA*) and the various species can have different lengths. Any species can be present in multiple copies, and the number of exemplars belonging to the various species will be denoted by $(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$. In this study the term "molecules" will be used to denote the exemplars belonging to the various species, either monomers or polymers. The reaction rates depend upon concentrations but, working at a fixed volume, we will sometimes refer to the number of molecules.

As in the original model proposed by Kauffman ([1986](#)), we consider two possible reactions, i.e., condensation and cleavage. Let us assume that both reactions need a catalyst to occur and that the kinetic of the spontaneous reactions is much slower than that of the catalyzed reactions, hence spontaneous reactions are neglected. We also assume that the rates of the reverse reactions are negligible with respect to that of the forward reactions. According to the number and the length of the species present in the system, the number of conceivable reactions is

$$\hat{R} = \sum_{i=1}^{N}(L(x_i) - 1) + N^2, \qquad (5)$$

where $L(x_i)$ is the length of the ith species and N is the total number of species present in the system (the sum refers to the number of cleavages and the quadratic term to the number of condensations). Note that in any particular simulation not all the possible reactions actually happen, and we assume that there is a probability p that a randomly chosen species catalyzes a randomly chosen reaction. Thus, in the present model there is no functional relationship between sequences (i.e. "chemical composition") of the catalysts and the reactions they catalyze.

Concentrations change in time according to an asynchronous stochastic approach based on the well-known Gillespie algorithm (Gillespie [1977](#), [2007](#)), by means of which at each step a reaction is chosen (among all the possible ones) and physical time is computed. All reactions occur in a well stirred chemostat and the concentration of each species is assumed to be constant all over the space. The model behavior thus depends on the characteristics of the initial chemical population and on the characteristics of the flux.

The dynamics

The initialization phase requires the creation of the initial population of species. It is possible to build all the species having a length up to the initial maximum species length M [the so-called *firing disk* (Farmer et al. [1986](#))], or to generate a more sparse initial population, for example by creating only N_{ini} species, i.e., a portion of the complete firing disk, according to particular distributions of length or composition. In the simulations described in this article all the species of length 1 are always created.

Once a species is created, its concentration is initialized according to some distribution having the length of the species as a parameter; so it is possible to generate initial compositions favoring short species or having a uniform distribution of species lengths. The total number R of conceivable reactions is computed and each species will catalyze each reaction independently with probability p (therefore, generating on average $r = R \cdot p$ catalyzed

reactions). The cutting point of cleavages is chosen with uniform probability and it is possible to tune the relative fraction of cleavages and condensations.

In condensation reactions two substrates are selected and the product is created concatenating them together. Note that, while cleavage is a bi-molecular reaction that occurs instantaneously when the catalyst binds the substrate, condensation is a three-molecular reaction and it is performed in two steps: in the first step the catalyst binds a substrate forming a temporary complex, while in the second step the complex binds the second substrate releasing the product and the catalyst. Note that also the spontaneous dissociation of the complex is considered. Hence, the system is composed of four possible kind of reactions (condensation is split in three phases):

Cleavage: $AB + C \rightarrow A + B + C$
Condensation: (whole reaction: $A + B + C \rightarrow AB + C$)

– Complex formation: $A + C \rightarrow A_C$
– Complex dissociation: $A_C \rightarrow A + C$
– Final condensation: $A_C + B \rightarrow AB + C$

The outgoing flux is simulated by means of a spontaneous decay of each specie and complex, with a common kinetic constant. The incoming flux is measured in moles per second.

The internal dynamics must take into account the fact that new species can be created; these new objects could react with the already present chemicals, and/or catalyze new reactions. So, whenever a new species is created, N increases, and the total number of conceivable reactions increases too; as a consequence reactions catalyzed by the species already present in the systems have to be updated according to the presence of the new species. If, in course of the dynamics, some of the existing species disappear, the current set of reactions is kept in memory, to assure consistency if those species would reappear.

Notice that the adoption of an asynchronous updating framework entails problems in identifying cycles. In order to analyze the structure of the system it is, in fact, useful to represent species and reactions by means of a directed graph, where the nodes are the species, and node X is linked to node Y if the species corresponding to node X catalyzes the formation of the specie corresponding to node Y. However, because of the asynchronous updating of the algorithm, there can be several network representations (a situation that is not encountered in the case of deterministic approaches, where all the reactions occur at the same time, corresponding to stable edges in the reaction graph representation).

Strictly speaking, on the graph representing the asynchronous framework at each step there is only one edge (representing the single reaction occurred during the last time interval). In order to avoid such a trivial representation, we decided to introduce a specific temporal window: the graph contains all the reactions which take place within the window. This representation is also useful to neglect the influence of very rare reactions, i.e., those with a very low rate but which might have occurred once in the past: if the reaction does not occur within the temporal window, the edge is removed from the graph. We define this graph as the "actual reaction graph".

Another possible representation is given by the graph containing all the possible reactions according to the species present at least in one copy at a given time t. This representation could provide indications about the "nearest adjacent possible" (Kauffman 2008) and is defined as "possible reaction graph".

Results of the analyses

In our model, an ACS is identified by a subgraph of the actual reactions graph in which each specie is formed by at least one reaction catalyzed by at least one other species present in such a subgraph; thus, each node belonging to an ACS is directly or indirectly connected with each other node of the ACS, i.e., it is a *strongly connected component* (*SCC* for short). This fact implies the presence of a path from a to b but also from b to a and, as a consequence, a *SCC* contains at least one cycle. If we define A as the adjacency matrix in which A_{ij} is equal to 1 when species j catalyzes the formation of species i and 0 if it does not, it is possible to determine whether an ACS is present by means of the analysis of the eigenvalue of A with largest real part (ELRP), λ_1. Since A is real and non-negative, the Perron–Frebonius theorem (Lutkepohl 1996) assures that λ_1 is real and $\lambda_1 \geqslant 0$; it can also be shown that if $\lambda_1 = 0$ there are no cycles in the graph, and the presence of a cycle is associated with $\lambda_1 \geqslant 1$ (Jain and Krishna 2001).

The variation of the residence time

A first set of simulations concerns the influence of the average residence time of the species present within the reactor on the overall dynamics, with particular regard to the emergence of ACSs.

Considering that the average residence time $\langle \psi \rangle$ is given by the following equation:

$$\langle \psi \rangle = 1/k_{out} \tag{6}$$

where k_{out} is the outflux rate, and that the average number of molecules present in the system at the equilibrium $\langle N_{eq} \rangle$ is given by the equation:

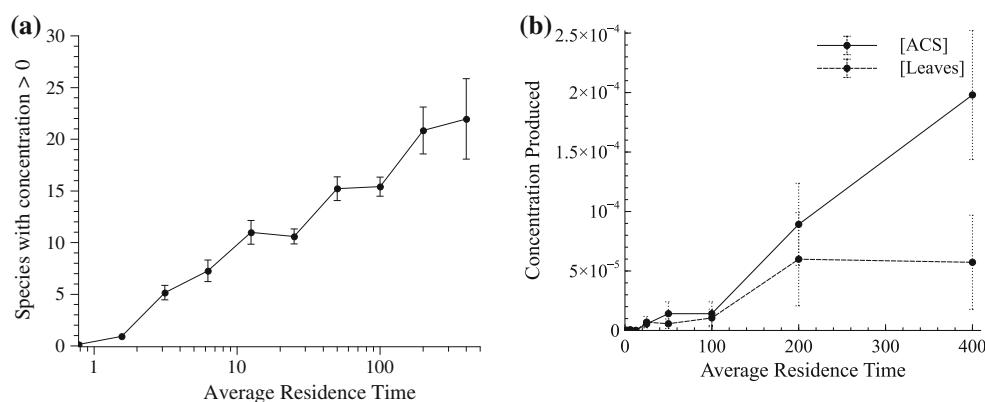$$\langle N_{eq} \rangle = I/k_{out} \tag{7}$$

Fig. 1 The graphs show: **a** The average number of species with concentration greater than 0 not belonging to the incoming flux in function of 10 different values of $\langle\psi\rangle$ (log scale on the *x*-axis), **b** The overall molecules concentration produced within the ACSs and by the

ACS first-order leaves in function of 10 different values of $\langle\psi\rangle$ (20 different runs, the *error bars* represent the standard error, $\langle\psi\rangle$: 0.78, 1.56, 3.13, 6.25, 12.5, 25, 50, 100, 200, 400 s)

where *I* is the number of molecules entering the reactor every second, we set some specific combination of influx and outflux to variate the average residence time, while keeping $\langle N_{eq}\rangle$ constant. Notice that, to sensibly analyze the emergence of ACSs we decided to set the reaction probability (i.e., the probability that a random species catalyzes a random reaction) close to the *critical* value, that is the value according to which every species catalyzes, on the average, one random reaction.[4] The results of the analyses are shown in Fig. 1.

In Fig. 1a, we can observe the dependence of the number of new species not belonging to the incoming flux at the end of the simulation in function of $\langle\psi\rangle$, which approximate a logarithmic curve.This first outcome suggests that larger residence times would entail a general enhancement of the activity of the system, which would favor the creation of new species and the maintenance of the existing ones.

In Fig. 1b, we can observe the total concentration produced within the ACSs an by the ACS leaves. Also in this case, the larger is the value of $\langle\psi\rangle$, the higher is the probability of encountering an ACS, highlighting the influence of this key parameter on the overall dynamical behavior.

The firing disk and the incoming flux

A second set of simulations is aimed to detect the possible effect of a variation of the composition of the firing disk, and relative concentrations, on the overall dynamics and on the emergence of ACSs. We test distinct initial configurations of the firing disks, in which we variate both the

Table 1 Average number of species with concentration greater than 0 not belonging to the incoming flux at the end of the simulation (30 different simulations for each table entry) starting from different initial conditions in terms of "the overall initial concentration" (conc) and "the number of different species present in the firing disk" (max length)

	Max length 2	Max length 3	Max length 4
Conc = 1e − 1	16.57	17.62	15.62
Conc = 1e − 2	17.23	15.76	15.97
Conc = 1e − 3	16.90	17.45	16.03
Conc = 1e − 4	18.00	17.14	17.72

maximum allowed length of the species (i.e., 2, 3, and 4) and their relative abundance in terms of initial concentration, while the influx composition and rate are kept constant (allowing all the species up to length 4 to enter the reactor). Also, in this case, the reaction probability is set critical and the average residence time is set to 50 s.

In Table 1, we can observe the variation of the average number of species with concentration greater than 0, not belonging to the incoming flux, present in the system at the end of the simulation (time = 1000 s). We can notice that no significant difference is detected when the composition of the firing disk is changed, pointing to a substantial independence of the overall dynamics from these parameters. Also the analysis about the emergence of ACSs confirms these results (not shown here).

We then performed further simulations in which different ensembles of networks are characterized by incoming fluxes composed of all the polymers up to length *Q*. The firing disks are kept equal to the corresponding influxes. It is important to remark that in this case, we allowed a supra-critical value of the reaction probability, to enhance the activity of the system and favor the emergence of ACSs. In this way the effects of a variation of the influx

---

[4] Notice also that in all the simulations that we describe in this article we use firing disks and incoming fluxes without any ACS and that in all the simulations the alphabet is composed of two letters, A and B.
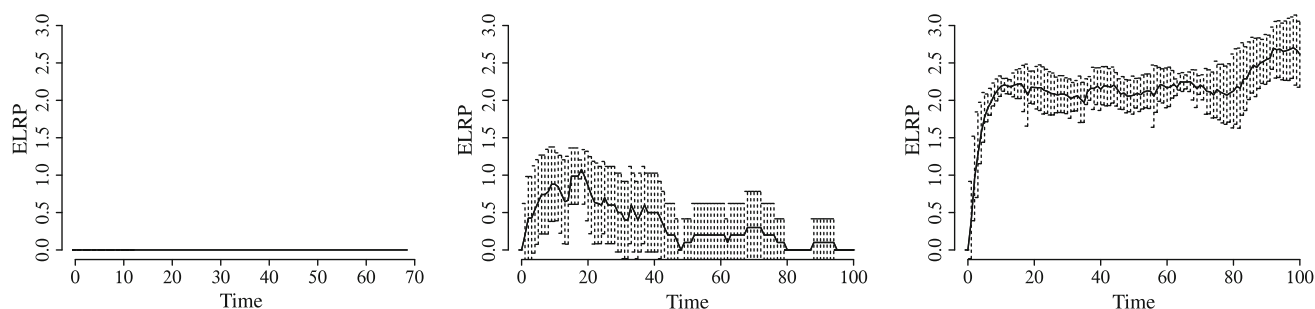
Fig. 2 Graphs show the average time behavior of the ELRP with respect to the heterogeneity of the influx (10 different runs, the *error bars* represent the standard deviation). From *left* to *right*: Influx composed by all the species up to length 2, all the species up to length 3 and all the species up to length 4
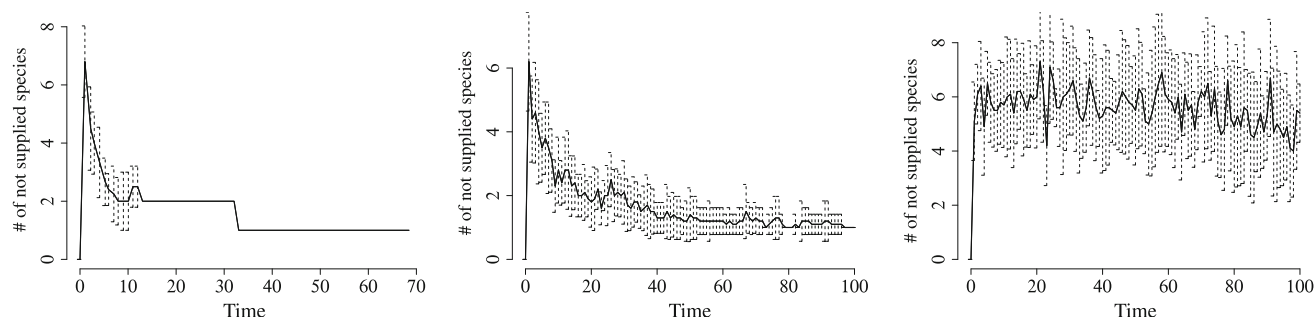

Fig. 3 Graphs show the average amount of catalysts not belonging to the influx with respect to the different compositions the influx (10 different runs, the *error bars* represent the standard deviation). From *left* to *right*: Influx composed by all the species up to length 2, all the species up to length 3 and all the species up to length 4

Fig. 4 Example simulation. On the *left panel* the number of molecules belonging to the ACS is shown whereas, on the *right panel*, the percentage of molecules (over the total number of molecules) belonging to an ACS versus time is shown. All simulations are made using a firing disk with maximum length 4
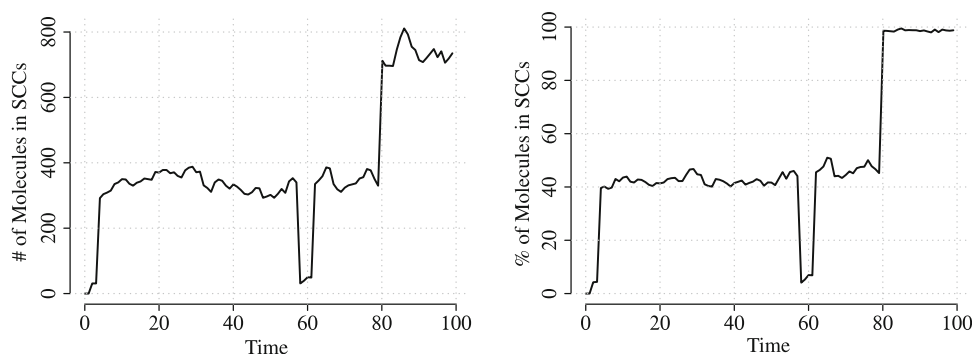
is more easily detectable. As one can see in Figs. 2 and 3, the system can generate and maintain ACSs only if Q is large enough.

The variation in time of the size of the ACSs in a typical simulation is shown in Fig. 4. We can observe that, despite their size, ACSs can undergo periods of strong instability (for example, close to second 60 during the simulation shown in Fig. 4), but also that in the long term the largest part of the molecules still present within the reactor are recruited by an ACS.

In order to prevent trivial results we forced the absence of ACSs within the incoming flux, therefore the existence of ACSs requires species which are generated in the reactor. Nevertheless, some of them are often present with a small number of molecules. This fact entails an important consequence: species characterized by a low number of molecules are subject to large fluctuations which can in turn affect also species present in greater number, but that are catalyzed by rarer ones.

The results can be summarized stating that ACSs may be subject to heavy random fluctuations and that simulations performed using a deterministic framework could not capture some important characteristics of the phenomenon. In practice, the ACSs that we could observe in the simulations appear to be structurally fragile, as depending on the presence of some rare molecules (or reactions), which
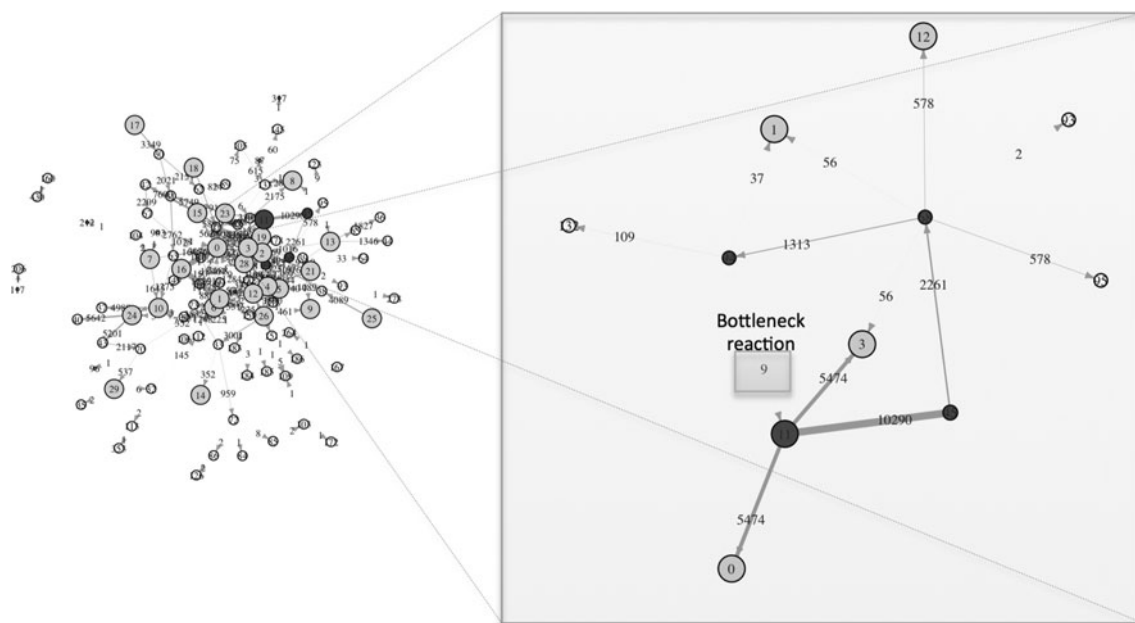
Fig. 5 A representation of a typical behavior of the system. A zoom of the reactions network (*right side*) shows four species (*dark grey nodes*) involved in an ACS characterized by a bottleneck reaction occurred only 9 times during the time span of the simulation. The other reactions involving the species forming the ACS occur 10290, 2261, and 1313 times. The bottleneck reaction occurs at a so slow rate that large temporal thresholds are needed to identify this structure as an ACS

are subject to random fluctuations and noise (Fig. 5). Moreover, given that this fragility already involves cycles composed by a very few components, we can reasonably hypothesize that the stability of bigger cycles (rarely observed) may degenerate, as occurs, for instance, with hypercycles (Eigen and Schuster 1977). Note also that this outcome could help in explaining some difficulties associated in observing the emergence of ACSs in the laboratory.

Conclusions and indications for further study

The model described in this study allows a proper treatment of the stochastic dynamics associated to catalytic reaction networks and therefore provides a virtual laboratory for investigating the influence of the key parameters and studying the general properties of the system.

Within this methodological framework it was possible to show that the composition of the incoming flux plays a major role on the dynamics, with particular regard to the emergence of ACSs, while, conversely, the variation in the compositions and concentrations of the firing disk appear to be substantially not relevant in the overall behavior. Furthermore, another key parameter of the system, the average residence time, appear to be largely correlated with an enhancement of the general activity, entailing a higher probability of emergence of ACSs when the time is higher.

The use of a stochastic approach is well grounded a priori, as we specified in the introduction, but the results of the simulations provide a further support to this choice. Indeed, one of the most interesting outcomes is the fact that, in most cases, the existence of the ACS critically depends upon that of some rare molecules, which may disappear because of fluctuations. This result provides a possible explanation of the observed difficulty in observing ACSs in wet lab experiments, even though further investigations are needed as, for instance, the introduction of the species forming the ACSs in the incoming flux, to look for possible differences in the behavior of these species in respect to those not belonging to any ACS.

Many developments of the model are possible as, for instance, the introduction of a relation between the structure (i.e. the chemical composition) of a catalyst and the kind of reactions that it will affect. Moreover, the model presented here has essentially an informational nature, while the energetics of chemical reactions are not explicitly considered. Since, however, such energetic effects play a major role in nature, a priority in future developments will be incorporation of such effects into the reactions within our model.

References

Bagley RJ, Farmer JD (1991) Spontaneous emergence of a metabolism. Artif Life II:93–140. Santa Fe Institute Studies in the Sciences of Complexity, X

Bower JM, Bolouri H (2004) Computational modeling of genetic and biochemical networks. Computational molecular biology. The MIT Press, Cambridge

Carletti T, Serra R, Poli I, Villani M, Filisetti A (2008) Sufficient conditions for emergent synchronization in protocell models. J Theor Biol 254(4):741–751

Dyson FJ (1985) Origins of life. Cambridge University Press, Cambridge

Eigen M, Schuster P (1977) The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. Die Naturwissenschaften 11(64):541–565

Eigen M, Shuster P (1979) The hypercycle: a principle of natural self-organization. Springer, Berlin

Farmer JD, Kauffman SA, Packard NH (1986) Autocatalytic replication of polymers. Physica D 22(2):50–67

Filisetti A, Serra R, Carletti T, Villani M, Poli I (2008) Synchronization phenomena in protocell models. Biophys Rev Lett 3(1/2):325–342

Filisetti A, Graudenzi A, Serra R, Villani M, De Lucrezia D, Fuchslin RM, Kauffman SA, Packard N, Poli I (2011) A stochastic model of the emergence of autocatalytic cycles. J Syst Chem 11:2. doi:10.1186/1759-2208-2-2

Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361

Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58(1):35–55

Jain S, Krishna S (2001) A model for the emergence of cooperation, interdependence, and structure in evolving networks. Proc Natl Acad Sci USA 98(2):543–547

Kaneko K (2006) Life: an introduction to complex systems biology. Understanding complex systems. Springer, New York

Kauffman SA (1986) Autocatalytic sets of proteins. J Theor Biol 119(1):1–24

Kauffman SA (2008) Reinventing the sacred: a new view of science, reason and religion. Basic Books, New York

Lutkepohl H (1996) Handbook of matrices. Wiley, New York, xvi + 304 pp, ISBN 0-471-96688-6

Munteanu A, Attolini C, Rasmussen S, Ziock H, Solé R (2006) Generic Darwinian selection in protocell assemblies. SFI-WP 06-09-032, SFI Working Papers, Santa Fe Institute

Rasmussen S, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF, Bedau MA (2004) Transitions from nonliving to living matter. Science 303:963–965

Segre D, Lancet D, Kedem O, Pilpel Y (1998) Graded autocatalysis replication domain (GARD): kinetic analysis of self-replication in mutually catalytic sets. Orig Life Evol Biosph 28(4–6):501–514

Serra R, Carletti T, Poli I (2007) Synchronization phenomena in surface-reaction models of protocells. Artif Life 13:1–16

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409(6818):387–390