Dynamical Stability of Autocatalytic Sets

Rudolf M. Füchslin[1,3], Alessandro Filisetti[1], Roberto Serra[2,1],
Marco Villani[2,1], Davide DeLucrezia[1], and Irene Poli[4,1]

[1] European Centre for Living Technology
Calle del Clero 2940, 30124 Venice, Italy
[2] Dipartimento di Scienze Sociali, Cognitive e Quantitative Universita di Modena e Reggio Emilia,
via Allegri 9, 42100 Reggio Emilia, Italy
[3] Artificial Intelligence Lab Univ. Zürich
Andreasstr. 15, CH-8050 Zürich, Switzerland
[4] Dipartimento di Statistica, Universita Ca' Foscari,
San Giobbe - Cannaregio 873, 30121 Venezia, Italy

Abstract

Theoretical investigations of autocatalytic sets rendered the
occurrence of self-sustaining sets of molecules to be a generic
property of random reaction networks. This stands in some
contrast to the experimental difficulty to actually find such
systems. In this work, we argue that the usual approach,
which is based on the study of static properties of reaction
graphs has to be complemented with a dynamic perspective
in order to avoid overestimation of the probability of getting
autocatalytic sets. Especially under the, from the experimen-
tal point of view, important flow reactor conditions, it is not
sufficient just to have a pathway generating a given type of
molecules. The respective process has also to happen with a
sufficient rate in order to compensate the outflow. Reaction
rates are therefore of crucial importance. Furthermore, pro-
cesses such as cleavage are on one hand advantageous for the
system, because they enhance the molecular variability and
therefore the potential for catalysis. On the other hand, cleav-
age may also act in an inhibiting manner by the destruction
of vital components: therefore, an optimal balance between
ligation and cleavage has to be found. If energy is included as
a limiting resource, the concentration profiles of the compo-
nents of autocatalytic sets are altered in a manner that renders
a certain range for the energy supply rate as optimal for the
realization of robust autocatalytic sets.

The results presented are based on a theoretical model and ob-
tained by numerical integration of systems of ODE. This lim-
its the number of involved molecular species which implies
that the quantitative findings of this work may have no direct
relevance for experimental situations, whereas the qualitative
insights in the dynamics of the systems under consideration
may generalize to systems of truly combinatorial size.

Keywords: Autocatalytic sets, autocatalytic metabolism, ori-
gin of life.

Introduction

In recent years, autocatalytic sets (ACS) Calvin (1956);
Eigen (1971) have attracted interest from many different re-
search directions. Probably most prominent are thereby in-
vestigations concerning the origin of life, but ACS proved to
be a concept also of value e.g. for the study of transitions
in general (non-chemical) systems of interacting production
processes including the generation of knowledge, see Hanel
et al. (2005).

Informally, the fundamental question with respect to
chemical reaction networks is whether or not a given set
of different, potentially catalytic molecules immersed into
a suitable environment (most often some type of flow reac-
tor) and provided with a sufficient supply of food or building
blocks is able of maintaining the concentration of its mem-
bers via mutual catalysis. The conditions under which such a
self-maintaining or autocatalytic set can be expected to ap-
pear with sufficiently high probability are then those to be
mimicked in an experiment e.g. concerned with the emer-
gence of protolife.

Based on different models of catalytic networks, there is
broad literature on the detection of ACS, see Letelier et al.
(2006); Mossel and Steel (2005); Hordijk and Steel (2004).
In Hordijk and Steel (2004) a polynomial-time algorithm
for the detection of an important class of ACS has been
presented. Hordijk and Steel applied this algorithm to a
model by Kauffman (1986). By analyzing large numbers
of randomly chosen networks, they corroborated a conclu-
sion which Kauffman derived from combinatorial reason-
ing, namely that in sufficiently diverse populations of po-
tentially catalytic chain molecules, an ACS will be present
almost with certainty. Thereby, ACS will form independent
of how sparse catalytic activity is distributed in the com-

binatorial variety of molecules, as long as this variety is big enough (usually limited by a maximal sequence length). Stated differently, given a certain variety of potentially catalytic molecules, there is always a threshold for the probability of catalytic activity such that above that threshold, ACS can be expected to emerge with high probability.

Despite some criticism (see Lifson (1997) and for a discussion of Lifson's arguments, see Steel (2000)) and the fact that more detailed models of catalysis may modify some results presented in Kauffman (1986), the main conclusions seem to generalize in one or the other form to a broad variety of models. The obvious question to ask then is, why ACS are not regularly discovered in the laboratory. In Filisetti et al. (2010), three possible answers were discussed. The first one (sometimes preferred by experimentalists) claims that the simplifications used in the formulation of the models on one hand make them tractable by analytical and/or computational means but on the other hand renders them unrealistic. The second answer (favored by some theorists) says that the basic statements derived from simplified models are also valid if the details of the physical and chemical world were considered, but that the threshold necessary for the emergence of ACS never has been reached. Finally, the third position (and also the one advocated in Filisetti et al. (2010) and in this work) highlights the fact that in investigations purely based on the properties of reaction graphs, dynamical and stochastic aspects are not considered. For some models, this is not necessary because their dynamics is basically (at least piecewise) determined by linear operators, e.g. Jain and Krishna (2001). But for most models (which are based on general reaction graphs), graph-theoretical methods may identify ACS which are only transient; this in the sense that the chemical dynamics eventually leads to a collapse of the ACS. This holds especially under flow reactor conditions, where e.g. a catalyst needs not only to be produced via some reaction path, but also at a sufficient rate in order to compensate for loss by outflow. Graph-theoretical means are able to identify whether or not a reaction path is present in a given network but not wether the dynamics establishes a non-trivial stationary ACS (In fact, one should speak of ACS exhibiting stationary or limit cycle behavior, but in practice one observes most models to yield almost exclusively stationary solutions. For a discussion, see e.g. Stadler et al. (1993)). In an experiment, however, it may be difficult to observe transient ACS, first because they may only be active during a very short period of time and second because their emergence may be highly susceptible to initial conditions. In contrast, stationary ACS which are able to produce a permanent deviation of some molecular concentrations from those one expects to result from the inflow and some non-catalytic background reactions offer a higher potential for being observable in a reproducible manner, as pointed out by Bagley and Farmer (1991). Whereas in Filisetti et al. (2010) the emphasis has been put on the investigation of the influence of stochastic fluctuations on the emergence and dynamics of ACS, this paper is concerned with the study of the influence of various parameters on the observability of stationary ACS.

The paper is organized as follows: In the second section, we discuss two different approaches for the definition of an ACS (or to be precise, the general and a more restrictive definition, the latter termed "autocatalytic metabolism") and motivate the choice being taken for the investigations in this work. In the third section, we briefly review the original model by Kauffman (1986) and present our implementation as a system of coupled ODEs. In the section reporting results, we show that the presence of a stationary ACS depends critically on the choice of parameters. We further study a derivative of the original model that takes energy considerations into account, means the different reactions compete for a, with a constant rate renewed, energy resource. We close with a discussion of the relevance of our results for experimental setups.

Autocatalytic Sets

We compare two different approaches for the analysis of autocatalytic sets. The first approach is especially appropriate for the study of reaction graphs and thoroughly discussed and formalized in Hordijk and Steel (2004). The second one, discussed in Bagley and Farmer (1991) takes into account the dynamics of the system but is less formal. Bagley and Farmer define an "autocatalytic metabolism" (ACM) as a coupled set of reactions which lead to permanent concentrations that are significantly departing from the values one would obtain without catalysis. As they point out, this definition is to some extent problematic, because what one regards as significant may depend on the experimental means. However, we will use a similar approach, because only those systems delivering a measurable deviation (both with respect to quantities as well as time) from some equilibrium distribution are of experimental interest. In order to highlight the difference between the two approaches, we briefly review the graph theoretical definition used by Hordijk and Steel and show that an ACS identified with their method needs not necessarily to be observable.

In Hordijk and Steel (2004) the main focus is laid on so called "reflexively autocatalytic and F-generated reaction systems (RAF)", whereby F denotes a set of "food"-molecules which are provided by the environment. For investigations concerned with the catalytic formation of chain molecules, F most often contains monomeric building blocks or a set of short oligomers. Informally, the concept of a RAF covers those sets of reaction systems R for which it holds that a) each reaction in R is catalyzed by a molecule being part of R and b) all reactants can be generated from a food set F by iterative applications of the reactions in R. In order to formalize the notion of a RAF in a rigorous manner, a number of definitions are required. We don't repeat

them here, but refer to the original work by Hordijk and Steel (2004)).

A RAF can be regarded as, once present, a potentially self-sustaining reaction system that in principle produces all the catalysts and intermediates it needs for its reactions. It is only potentially self-sustaining, because necessary molecules need not only to be produced but being produced with sufficient rates. Note further that the definition of a RAF does not require the system to emerge, given the molecules in F are supplied (In fact, the elements of F need not to be catalysts at all).

As shown in Hordijk and Steel (2004), there exists a polynomial-time algorithm for the detection of RAFs, given a system of catalytic reactions. That such a RAF is only potentially self-sustaining is demonstrated by a (completely artificial) reaction system given as follows (with respective catalyst and reaction rate above the arrows):

$$a + e \xrightarrow{d,k} c$$
$$b + e \xrightarrow{c,k} d$$
$$c \xrightarrow{d,k} e + e$$

$$(1)$$
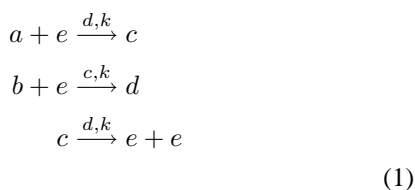
With $F = \{a, b\}$, this system qualifies as a RAF (possibly being part of some bigger catalytic reaction system). It is possible (not shown here) to add further reactions representing the renewal of resources and outflow, the former taking place with unit rate, the latter with rate k_d. Setting $k = 1$ and $a(0) = b(0) = c(0) = d(0) = e(0) = 1$, the behavior of the system then depends critically on the size of k_d. As illustrated in Fig. 1, the system attains a stationary state for $k_d = 0.1$ and collapses for $k_d = 0.5$. This observation is of importance insofar that it shows that one tends to overestimate the probability for the observation of experimentally relevant ACM if one relies on static, graph theoretical methods yielding probabilities for the occurrence of ACS. Consequently, in what follows we employ dynamic reaction kinetics in order to decide whether a reaction system contains as a subsystem an ACM in the sense of Bagley and Farmer (1991).

The Model

A fundamental model for the study of the emergence of ACS has been proposed in Kauffman (1986); we will briefly review this approach and its main conclusions and present our own implementation which is used for the construction of a set of ODEs. These ODEs are solved numerically for various parameter settings in order to identify the relative importance of different reaction mechanisms. Thereby, we are interested in parameter combinations that exhibit non-trivial optima for the probability of the existence of an ACM, especially if these parameters offers the potential of being controllable in an experimental setting.
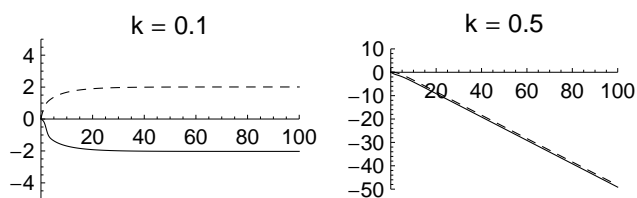

Figure 1: Time evolution of the system given by eqs. 1 for two different values of the outflow rate parameter k_d. Shown are the logarithms of the concentration of $c(t)$ (continuous line) and $d(t)$ (dashed line) as a function of time.

The Basic Model

In Kauffman (1986), the properties of sets of potentially catalytic di-block copolymers were investigated. Thereby, it was assumed

- Polymers consist of two different types of monomers A and B.

- There are two types of catalyzed reactions, namely ligation and cleavage.

- The probability for a polymer P_c to catalyze a ligation $P_1 + P_2 \xrightarrow{P_c} P_1 P_2$ or a cleavage $P_1 P_2 \xrightarrow{P_c} P_1 + P_2$ is given by a probability r.

- The number p_i represents the density of the polymer P_i.

This setting, basically a random reaction system, doesn't make any specific "helpful" assumptions supporting the emergence or existence of an ACM, and nevertheless, strong evidence was given that such a system should eventually contain an ACM, given only a sufficiently large variety of different polymers being included in the system (In case of block polymers, this can be achieved simply by allowing sequences of length up to a critical L_c).

Several implementations of random graph models using ODEs have been studied, see e.g. Farmer et al. (1986); Bagley and Farmer (1991). In this work, the dynamics of

the system is given by:

$$\frac{dp_i}{dt} = k_{i,in} - k_{out}p_i \qquad (2)$$

$$+ \sum_{j,k,m} k_{j,k,L}L(j,k,i,m)p_jp_kp_m$$

$$- \sum_{j,k,m} k_{i,j,L}L(i,j,k,m)p_ip_jp_m$$

$$- \sum_{j,k,m} k_{j,i,L}L(j,i,k,m)p_jp_ip_m$$

$$+ k_C \sum_{j,k,m} C(i,j,k,m)p_kp_m$$

$$+ k_C \sum_{j,k,m} C(j,i,k,m)p_kp_m$$

$$- k_C \sum_{j,k,m} C(j,k,i,m)p_ip_m.$$

Thereby, p_i represents the density of a polymer with sequence P_i composed of two types of monomers A, B. The rate of influx $k_{i,in}$ is set to one for the monomers A, B and zero for all other sequences. Outflow is determined by the rate k_{out}, and the kinetic rates of ligation and cleavage are denoted by $k_{i,j,L}$ and k_C respectively. The arrays L and C represent the random graphs, chosen at the beginning of each run: This means that L, C are arrays representing fixed random reaction networks, which, once set, remain constant. Using the symbol $\oplus$ for sequence concatenation, it holds:

$$L(i,j,k,m) = \begin{cases} 0 & P_i \oplus P_j \neq P_k \\ 1 \text{ with prob. } \quad r_L & P_i \oplus P_j = P_k \end{cases} \qquad (3)$$

and

$$C(i,j,k,m) = \begin{cases} 0 & P_i \oplus P_j \neq P_k \\ 1 \text{ with prob. } \quad r_C & P_i \oplus P_j = P_k \end{cases} \qquad (4)$$

The index m represents the dependence on the catalyst P_m.

In all calculations subsequently shown, several additional assumptions have been made:

1. The monomers A, B must not act as catalysts; this in order to enhance chemical plausibility.

2. There is a maximal sequence length L. Ligations may well produce longer sequences, but those are assumed to fall out by precipitation. This is physically plausible and keeps the system tractable.

3. In order to capture steric effects, the ligation rate $k_{i,j,L}$ is length dependent. Shall $|P_i|$ denote the length of P_i, we set $k_{i,j,L} = k_L/(|P_i||P_j|)$ for some constant k_L. The idea behind this (crude) approximation is that in a well-stirred reactor, the collision frequency of two sequences is assumed to be independent of the length. The collision

happens by the contact of two monomers, one out of each sequence. The chance that those are the ones that are able of mutual ligation because they mark the end and the start of the respective sequences is inversely proportional to the respective length of the sequences.

The system then contains $2^{L+1} - 2$ variables. This means, taking into account the non-catalycity of the monomers, that there are $(2^{L+1} - 2)^2(2^{L+1} - 4)$ potential ligation reactions and $(2^{L+1} - 4)\sum_{l=2}^{L} 2^l(l-1)$ possible cleavage processes. As it turned out, already values of $L = 6$ deliver systems of sufficient combinatorial variety in order to exhibit interesting dynamical effects. In all simulations, we set $\forall i : p_i(0) = 1$ as initial condition; this with the idea to give a potential ACM in a random graph sufficiently favorable starting conditions. Following Bagley and Farmer (1991), a random reaction graph qualifies as containing an ACM, if the concentration of at least one non-monomeric species is above a threshold T after a time interval longer than $10t_d$ with $t_d = -\log(T)/k_{out}$ denoting the typical decay time for T. As will be shown (and has already been discussed by Bagley and Farmer), the decision whether a reaction system contains an ACM is surprisingly insensitive to the choice of T. The numerical solutions were obtained by internal routines of the software package Mathematical^{TM} and a sample of solutions was verified with a standard adaptive fourth-order Runge-Kutta solver.

The Model with Explicit Consideration of Energy

Most of the investigations dealing with ACM don't take into account energy considerations, or more generally, the explicit competition for some limited resource other than the supplied monomers. As will be discussed in the result section, such an external limitation need not to be disadvantageous for the system, but may even help to stabilize it. We consider energy in a relatively simple manner. The ligation and cleavage terms in eqs. 2 are multiplied with the concentration $e(t)$ of some energy resource. Thereby, the energy resource is used up and permanently renewed by inflow with a rate k_E. The dynamics of the additional variable $e(t)$ is given by:

$$\frac{de}{dt} = k_E - k_{out}e \qquad (5)$$

$$- \sum_{i,j,k,m} k_{i,j,L}L(p_i,p_j,p_k,p_m)p_ip_jp_me$$

$$- k_C \sum_{i,j,k,m} C(p_i,p_j,p_k,p_m)p_kp_me.$$

Results

In this section, we study the dependence of the dynamics of the models presented in the preceding section. Some of the parameters remain fixed for all simulations: $k_{out} = 0.02, k_L = k_C = 1$. Furthermore, each data point representing an average value has been computed using at least 20,
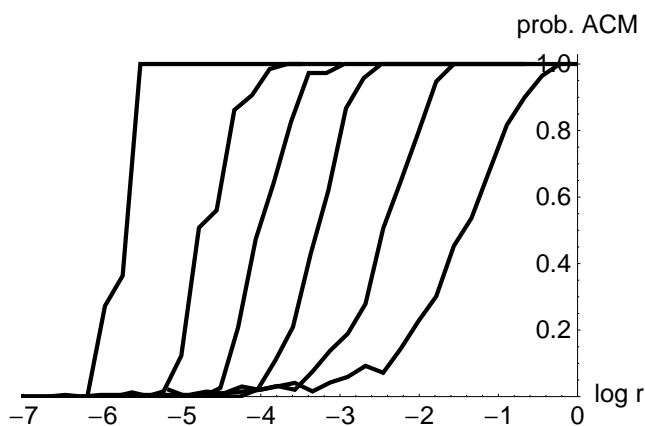
Figure 2: Probability for observing an ACM in a random reaction graph as a function of the catalytic reaction probability $r_L = r_C = r$ for different values of the maximal sequence length $L = 2, 3, 4, 5, 6, 8$. Starting from $L = 8$, graphs representing decreasing length exhibit increasing values for the transition value of r.

Figure 3: Average size of ACM (number of non-monomeric components bigger than $T = 10^{-6}$ after $t = 10^5$) as a function of r and for sequence length $L = 3, 4, 5, 6$ (bottom to top). Shown are the median values for the size of the ACM, the whiskers denoting the first and third quartile. Above the transition value of r, the system tends to be maximally diverse (A maximal sequence length L implies $2^{(}L + 1) - 4$ non-monomeric sequences).

but most often more than 50 samples. As a convention, logarithms are always taken to the base e. Whiskers, if shown, denote first and third quartiles.

The Fundamental Transition

As postulated in Kauffman (1986), for sufficiently large values of the probabilities for catalytic reactions r_L and r_C given in eqs. 3 and 4, the reaction graph should contain an ACM with high probability. In fig. 2, this transition is clearly observable and becomes sharper for longer sequences. Interestingly, the transition curves, giving the probability of observing at least one non-monomeric sequence with a concentration above the threshold value T look identically the same for T in the range from 10^{-12} to 10^{-2}, which means that if there is an ACM, at least one of its components will be present with a significant concentration. Fig. 3 shows the average size of the ACM, means the average number of components with concentration values above a threshold $T = 10^{-6}$ after an integration time $t = 10^5$ for sequences of maximal length $L = 3, 4, 5, 6$. We observe that above the transition value of r, the system becomes maximally diverse. This may be of relevance in an evolutionary context.

The Role of Cleavage

Given a certain fixed probability for ligation r_L, one may ask for the corresponding optimal value of r_C. It is clear that cleavage has some beneficial aspects for the appearance of an ACM, because cleavage tends to enlarge the variety of sequences. However, cleavage may as well destroy vital parts of an ACM. This is relevant especially under flow reactor conditions, where the generation of a specific sequence
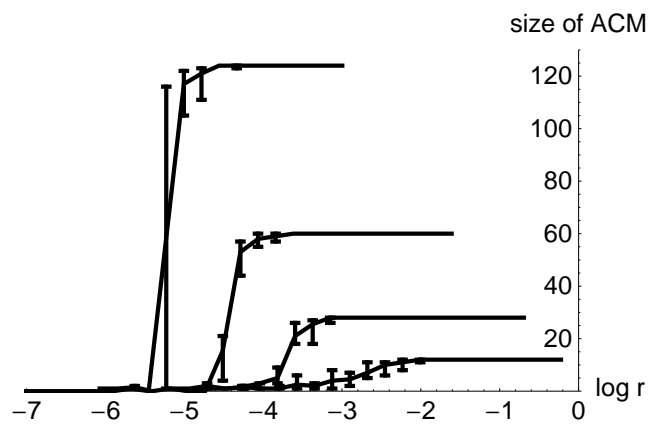
needs to be sufficiently powerful in order to compensate the outflow. And in fact, in fig. 4, a clear optimum for r_C can be observed, given a fixed $r_L = 0.01$ and $L = 6$. Notably, in our simulation, this optimum perfectly justifies the original choice of $r_L = r_C$ by Kauffman. The choice of r_L in the transition region is motivated by first taking into account that a system may be based only on ligation but not solely on cleavage (at least with monomeric input). A small value for r_L will most probably not yield an ACM. A large value is also not of big interest: A system with lots of ligation reactions already produces most sequences and does not profit from a further broadening of the sequence variety by cleavage. The transition region in fig. 2 is the domain in which an optimization of r_C will take the most effect.

Again, it is emphasized that the curve shown does not depend on the detection threshold T, though the average number of concentrations above the threshold does, see figs. 5 and 6. Note that whereas the curve in fig. 4 refers to the whole sample and shows the ratio of those reaction systems containing an ACM, the data in figs. 5 and 6 give the average size of the ACM, provided there is one. Consequently, data points at the lower and higher end of the scale are of less statistical weight (and relevance) than those in the middle.

The Role of Energy

Controlling the influx of energy (or, to be chemically more accurate, the influx of molecular energy carriers) is a parameter easy to control in an experiment, therefore its influence is of interest. It is clear that below a certain threshold of the influx rate k_E the generation of non-monomeric
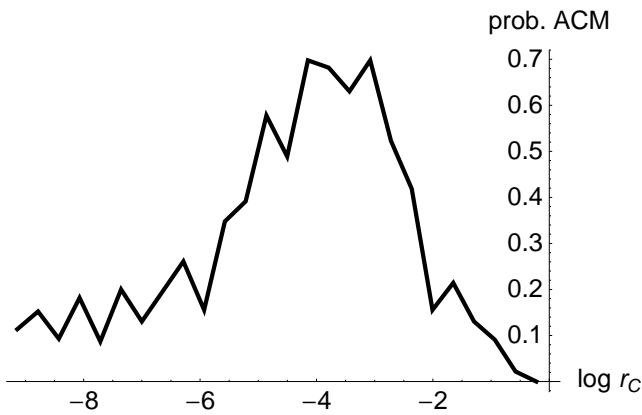
Figure 4: Probability for observing an ACM in a reaction graph with maximal sequence length $L = 6$ and $r_L = 0.01$ as a function of r_C. The detection threshold is set to $T = 10^{-6}$ (continuous line) and $T = 10^{-2}$ (dashed line).
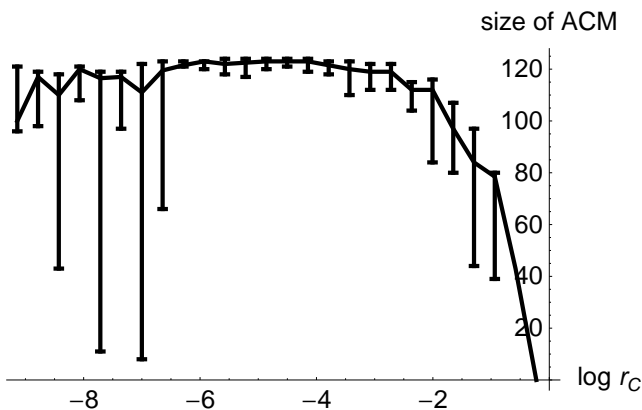

Figure 5: Average size of ACM for $L = 6$ and $r_L = 0.01$ as a function of r_C. The detection threshold is given by $T = 10^{-6}$. Shown are the median values for the size of the ACM and the whiskers denote the first and third quartile.
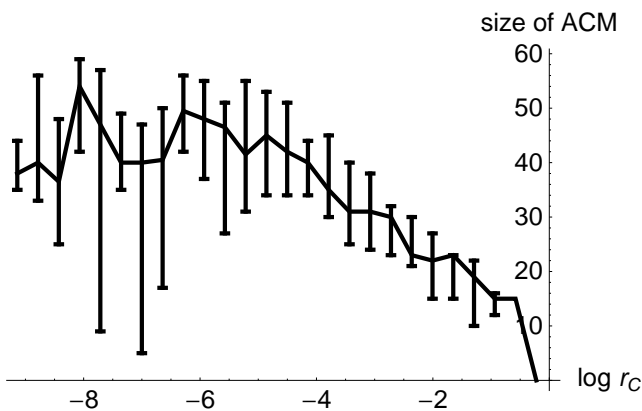

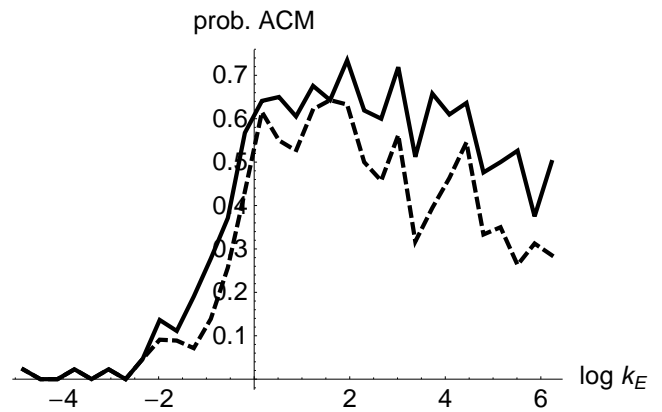Figure 6: Same as fig. 5, but with $T = 10^{-2}$.

Figure 7: Probability for observing an ACM in a random reaction system with $L = 6, r_L = r_C = 0.01$ as a function of the rate of energy influx k_E.

sequences is not anymore powerful enough to compensate for the outflux. This can be seen in fig. 7. Given suitable system parameters, ACM are easy to observe at higher values of k_E. Interestingly, the average size of the ACM for a large threshold T shows a maximum for intermediate values of k_E, see fig. 8 (giving the average number of concentrations above $T = 10^{-6}$) and more prominently for $T = 10^{-2}$ in fig. 9. A possible explanation for this phenomenon is that the plenty abundance of energy alllows the generation of more or less all possible sequences, as suggested by the results shown in fig. 3. A more fierce competition for energy, however, may lead to the eventual extinction of some side branches of an ACM and consequently a boost of its "core" components. This externally controlled focussing is of relevance, because in more realistic scenarios with larger sequence lengths, the relative concentrations of core components may be much lower than in the (numerically tractable) model systems presented in this work. Consequently, stochastic fluctuations play a more important role and a mechanism strengthening the "backbone" of an ACM at the expense of some side reactions increases the robustness of the system which is of evolutionary and experimental importance (the consideration made here applies also to the scenario discussed in fig. 6). Studying stochastic effects in ACM with longer sequences requires, however, a particle based approach. For a detailed discussion, see Filisetti et al. (2010).

Summary and Discussion

We have shown the importance of the dynamics of a reaction system for answering the question whether it contains an autocatalytic metabolism. Many algorithms are based on the analysis of combinatorial properties of random graphs. Thereby, they are not considering that, especially in the situation of a flow reactor, there must not only be a pathway
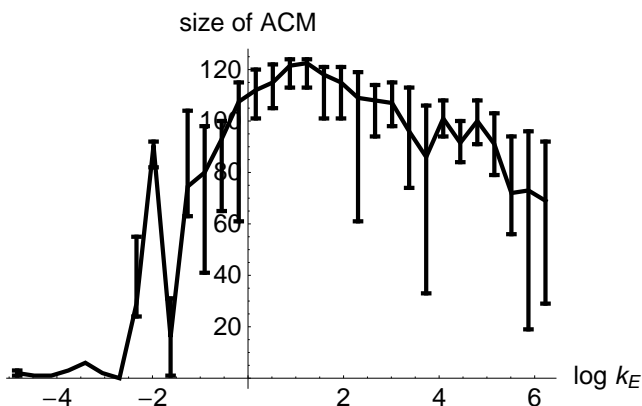
Figure 8: Average size of the observed ACM in a random reaction system with $L = 6, r_L = r_C = 0.01$ as a function of the rate of energy influx k_E and for a detection threshold $T = 10^{-6}$.
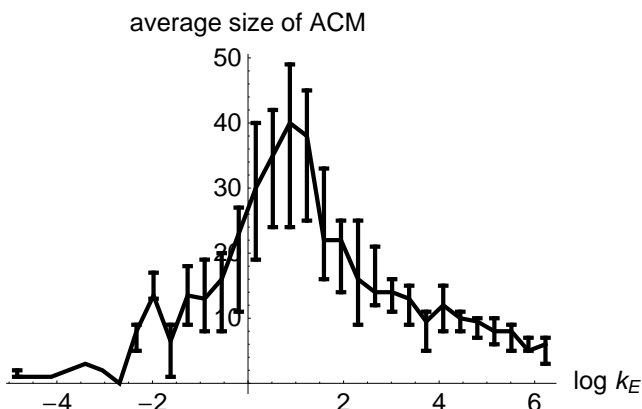

Figure 9: Average size of the observed ACM in a random reaction system with $L = 6, r_L = r_C = 0.01$ as a function of the rate of energy influx k_E and for a detection threshold $T = 10^{-2}$.

for the production of a given molecule but its production has in addition to happen at a rate that compensates for the loss by outflow. Studying the kinetic behavior of random reaction systems reveals the importance of a proper balancing of the probabilities for different types of reactions: We investigated cleavage and found that taking into account dynamics, cleavage does not only enlarge the variety of polymer species (which is desirable from the perspective of obtaining an ACM) but may also destroy components relevant for the system with a rate that cannot be compensated by their respective generation processes. We also investigated the role of energy consumption and found that the introduction of energy as a limiting factor strongly influences the concentration profile of the ACM. It turned out that whereas a large supply of energy leads to a broad variability of sequences, intermediate values seem to favor ACM with less, but, with respect to concentration also in absolute terms, more pronounced components. This means that such intermeidate values render ACM that are less susceptible to fluctuations, which is of relevance in the context of evlutionary processes.

We investigated systems with rather short sequences, mostly with a maximal sequence length of $L = 6$. The numerical values for the catalytic probabilities r_L and r_C need then to be of a size which is chemically not realistic. We claim that our results are of worth because whereas the quantitative features of the shown results heavily depend on L, the qualitative ones don't. Even more, data (partially not shown) suggests that the discussed effects become more pronounced with increasing L. According investigations need then to be performed in a particle based manner, see Filisetti et al. (2010). Another interesting perspective is presently investigated by DeLucrezia and coworkers. In their approach, the "monomers" are replaced by pre-prepared strands consisting of some ten amino acids. A sequence consisting of a combinatorial assembly of these strands may have a higher probability of exhibiting catalytic properties. However, the model presented in this paper is then only a "coarse-grained" approximation to the dynamics, because cleavage may well happen within one of the original monomeric strands.

Our choice of the initial conditions, namely to set the concentrations of all sequences to one at the start is certainly unrealistic and motivated by our focus on stability considerations. The discovery that the energy supply influences the concentration profile opens the perspective of "iterative" emergence. A very limited set of initially provided components may establish a first, still frail ACM which produces as side products some further, possibly catalytic components at low concentrations. A only temporal increase of the energy supply may enable the system to reach a new basin of attraction by a short-term increase of cleaving activity which in turn produce a passing wider variety of sequences at sufficient concentration in order to take effect, but without having to cope with the long-term presence of enhanced cleavage. We will address this scenario in a subsequent work fo-

cussed on issues of emergence, also considering aspects of stabilization against molecular parasites achieved by spatial organization with Filisetti et al. (2008) or without Füchslin and McCaskill (2001); Füchslin et al. (2004) explicit compartmentalization.

The problem of deciding whether or not a given reaction system contains an ACM may one remind to a similar problem in systems biology, namely the determination of possible fluxes in a only partially known metabolic networksVarma and Palsson (1994); Orth et al. (2010). In flux balance analysis, one basically determines the set of potential solutions for the fluxes, given that a) the stoichiometric matrix and a vector containing fluxes forms an underdetermined linear system and b) some (in practice usually linear) constraints have to be observed. Flux balance analysis provides a highly successful and efficient tool for e.g. the optimization of only partially known networks (By using linear programming). The problem we address in this work is, however, different. The networks are completely known and therefore, the flux balance equation are fully determined, which means that searching a stationary solution requires solving a non-linear system.

Taking into account dynamics shows that first, one of the reasons for the fact that spontaneously formed autocatalytic systems have not or only rarely been observed in the laboratory may not only be due to lack of catalytic activity. As a matter of fact, it could even be caused by too much catalysis, if cleavage is too frequent. Second, and probably more important, we need to shift our attention from focussing solely on catalysis (and respective probabilities) to a picture in which kinetics plays an important role too. Even if we had reaction system in which in principle an ACM could produce measurable signals, it only does if the kinetic parameters are suitably chosen. Some of these parameters, such as e.g. outflux rates, can easily be manipulated in an experiment and should be in the focus of future work.

Acknowledgments

References

Bagley, R. and Farmer, J. (1991). Spontaneous emergence of a metabolism. *Artifcial Life II*, pages 93–140. Santa Fe Institute Studies in the Sciences of Complexity X.

Calvin, M. (1956). Chemical evolution and the origin of life. *Amer. Sci.*, 44:248–263.

Eigen, M. (1971). Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58:465–523.

Farmer, J. D., Kauffman, S. A., and Packard, N. H. (1986). Autocatalytic replication of polymers. *Physica D*, 22(2):50–67.

Filisetti, A., Serra, R., Carletti, T., Villani, M., and Poli, I. (2008). Synchronization phenomena in protocell models. *Biophysical Reviews and Letters*, 3(1/2):325–342.

Filisetti, A., Serra, R., Villani, M., Füchslin, R. M., Packard, N., Kauffman, S. A., and Poli, I. (2010). A stochastic model of autocatalytic reaction networks. submitted to ECCS10.

Füchslin, R. M., Altmeyer, S., and McCaskill, J. (2004). Evolutionary stabilization of generous replicases by complex formation. *Europ. Phys J. B*, 38(1):103–110.

Füchslin, R. M. and McCaskill, J. S. (2001). Evolutionary self-organization of cell-free genetic coding. *Proceedings of the National Academy of Science USA*, 98(16):9185–9190.

Hanel, R., Kauffman, S. A., and Thurner, S. (2005). Phase transition in random catalytic networks. *Physical Review E*, 72:036117.

Hordijk, W. and Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227:451–461.

Jain, S. and Krishna, S. (2001). A model for the emergence of cooperation, interdependence, and structure in evolving networks. *Proceedings of the National Academy of Science USA*, 98(2):543–547.

Kauffman, S. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119:1–24.

Letelier, J.-C., Soto-Andrade, J., Abarzua, F. G., Cornish-Bowden, A., and Cardenas, M. L. (2006). Organizational invariance and metabolic closure: Analysis in terms of (m;r) systems. *Journal of Theoretical Biology*, 238:949–961.

Lifson, S. (1997). On the crucial stages in the origin of animate matter. *J. Molecular Evolution*, 44:1–8.

Mossel, E. and Steel, M. (2005). Random biochemical networks: the probability of self-sustaining autocatalysis. *Journal of Theoretical Biology*, 233:327–336.

Orth, J. D., Thiele, I., and Palsson, B. O. (2010). What is flux balance analysis? *Nature Biotechnology*, 28:245–248.

Stadler, P. F., Fontana, W., and Miller, J. H. (1993). *Physica D*, 63:378.

Steel, M. (2000). The emergence of a self-catalysing structure in abstract origin-of-life models. *Appl. Math. Lett.*, 3:91–95.

Varma, A. and Palsson, B. O. (1994). Flux balancing: Basic concepts, scientific and practical use. *Biotechnology*, 12:994–998.