SPATIAL MODELS OF PREBIOTIC EVOLUTION: SOUP BEFORE PIZZA?

ISTVÁN SCHEURING[1]*, TAMÁS CZÁRÁN[1]*, PÉTER SZABÓ[2], GYÖRGY KÁROLYI[3] and ZOLTÁN TOROCZKAI[4]

[1] *Department of Plant Taxonomy and Ecology, Research Group of Ecology and Theoretical Biology, Eötvös University, Pázmány P. sétány 1/c, H-1117 Budapest, Hungary;* [2] *Adaptive Dynamics Network, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria;* [3] *Program in Applied Mathematics, University of Arizona, 617 N Santa Rita, 85721 Tucson, Arizona, USA*

Department of Structural Mechanics and Center for Applied Mathematics and Computational Physics, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1521 Budapest, Hungary; [4] *Complex Systems Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, 87545, USA*

(author for correspondence, e-mail: shieazsf@ludens.elte.hu)*

(Received 17 July 2002; accepted in revised form 9 September 2002)

Abstract. The problem of information integration and resistance to the invasion of parasitic mutants in prebiotic replicator systems is a notorious issue of research on the origin of life. Almost all theoretical studies published so far have demonstrated that some kind of spatial structure is indispensable for the persistence and/or the parasite resistance of any feasible replicator system. Based on a detailed critical survey of spatial models on prebiotic information integration, we suggest a possible scenario for replicator system evolution leading to the emergence of the first protocells capable of independent life. We show that even the spatial versions of the hypercycle model are vulnerable to selfish parasites in heterogeneous habitats. Contrary, the metabolic system remains persistent and coexistent with its parasites both on heterogeneous surfaces and in chaotically mixing flowing media. Persistent metabolic parasites can be converted to metabolic cooperators, or they can gradually obtain replicase activity. Our simulations show that, once replicase activity emerged, a gradual and simultaneous evolutionary improvement of replicase functionality (speed and fidelity) and template efficiency is possible only on a surface that constrains the mobility of macromolecule replicators. Based on the results of the models reviewed, we suggest that open chaotic flows ('soup') and surface dynamics ('pizza') both played key roles in the sequence of evolutionary events ultimately concluding in the appearance of the first living cell on Earth.

Keywords: coexistence, early evolution, Eigen's paradox, prebiotic soup, open chaotic flow, prebiotic pizza, parasites, RNA-world

1. Introduction

Origin of life is a comprehensive research subject with many unsolved problems ranging from the prebiotic synthesis of biologically important organic molecules (like ribose and pyrimidine bases etc.) to the emergence of the first proto-cells including autocatalytic metabolic cycles and replicating information carrier mac-

Origins of Life and Evolution of the Biosphere **33:** 319–355, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*

romolecules (Maynard Smith and Szathmáry, 1995). It seems highly probable that these proto-cells were preceded by 'naked' self-replicating macromolecules (Maynard Smith and Szathmáry, 1995) living in the 'primordial soup' or on the surface of 'prebiotic pizza'. Since recent RNAs are capable of replication (with some catalytic help), and they are also shown to have catalytic activity themselves (Zaug and Cech, 1986; Doudna and Cech, 2002), the assumption that this intermediate stage of prebiotic evolution might have been dominated by self-replicating, RNA-like nucleic acid oligomers seems to be a plausible one (Gilbert, 1986; Joyce and Orgel, 1999), but there are other opinions (Shapiro, 1984; Kauffman, 1986; Segré *et al.*, 2001). In the ancient RNA world RNA-like oligomers must have been able to replicate without the aid of specific peptide enzymes, and they must have competed for a few limiting resources such as mononucleotids and other energy rich compounds.

Replication has a small chance of being perfect: due to mononucleotide pairing mismatches the copy is rarely identical to the original template. Replication is even less accurate without specific replicase and proofreading enzymes, which were definitely not present in the ancient RNA world. In his keystone paper Eigen (1971) suggested a mathematical model describing the mutation-selection process of these ancient replicating macromolecules. The most important result of his study was that the maximum amount of selectively maintained information, i.e., the maximum length of the fittest replicator molecule, is limited by copying fidelity. More precisely, he showed that $N < \ln(s)/(1-q)$, where N is the maximum number of nucleic acid monomers in the replicator molecule containing biologically meaningful information, s is the selective superiority of the fittest replicator (i.e., the replication rate of the fastest replicator divided by that of the average of all the rest) and q stands for the per nucleotide per replication copying fidelity. If the fittest macromolecule is longer than this limit, its concentration will be very low, and it can easily disappear from the system by stochastic drift (Szathmáry, 1989). Estimating the selective superiority s of the best replicator to be roughly between 10 and 100, and the copying accuracy q per nucleotide without replicase enzymes between 0.8 and 0.9 (Ekland and Bartel, 1996; Johnston *et al.*, 2001), it is concluded that the maximum length of persistent molecules ranges from 10 to 100 nucleotides (Eigen, 1971; Maynard Smith and Szathmáry, 1995; Scheuring, 2000). This means that the winner of the mutation-selection process is the fittest macromolecule consisting of some dozens of nucleotides, and the system (so called *quasispecies*) also maintains a distribution of the closest mutants of the winner type (Eigen, 1971; Eigen and Schuster, 1979). Given that even the simplest RNA viruses – incapable of independent self-replication – carry a genome longer than 4K monomers, there is a considerable gap between the first – hypothetical – replicators and the genomes of the first protocells. Surmounting this gap requires specific replicase enzymes which could increase copying fidelity and thus the length of persistent replicators. Non-enzymatic replication can maintain short replicators only, but these are insufficient for coding replicases. This is 'Eigen's paradox' or the

'Catch 22' of prebiotic evolution (Maynard Smith, 1983; Szathmáry, 1989): no genome without enzymes, and no enzyme without genomes.

One way out of this trap is through maintaining the coexistence of several different replicator molecules, so that the information necessary for coding a replicase enzyme can be stored and transmitted by a 'community' of smaller information carriers. Since the coexistence of different species is a typical ecological problem, we call these approaches of Eigen's paradox the *ecological solution*.

There is an alternative resolution to the problem which assumes that the template and the enzymatic functions have *coevolved*: the length of the replicators and their replicase activity, copying fidelity have increased in parallel (Poole *et al.*, 1999; Scheuring, 2000; Szabó *et al.*, 2002).

Naturally, all these scenarios are feasible only if these replicator systems are resistant against invasion by any possible mutant replicators. Unfortunately, *neither* of the suggested model systems are resistant against such parasitic selfish replicators, if the replicating molecules are supposed to be *point-like* objects in a *perfectly mixed* medium. (That is, ordinary differential equation models are vulnerable to parasite invasion.) Therefore, the common feature of the different models we consider below is that they suppose macromolecules to be *discrete* entities interacting with each other in a local manner, and moving by diffusion on a surface or advected in a flow of aquatic medium. Spatially extended modeling is undoubtedly a more adequate description of reality than most of the classical analytical models, and, interestingly, it concludes in resistance against selfish parasitic replicators as well.

Below the most important spatially extended models are discussed in which Eigen's paradox is resolved either by the coexistence of different replicators or by their coevolution. After focusing on the advantages and drawbacks of different scenarios, a coherent mosaic from the existing tiles is constructed. In Section 2, coexistence of replicators is shown to be ultimately related to non-trivial spatial distribution and to the discrete size of the competing replicators. First the metabolic network and hypercycle model are reviewed when individuals are adsorbed on a prebiotic surface. Then it is shown that replicators being purely competing for limiting resources or being in metabolic coupling can coexist in open chaotic flows. In Section 3, the coevolutionary scenario of information integration is presented. Finally, in the Discussion we suggest possible elementary transitions leading from naked RNA-like replicators to the first proto-cells.

2. Coexistence of Replicators

2.1. HYPERCYCLES ON A SURFACE

The first ecological solution to Eigen's paradox was the hypercycle: a system of autocatalytic replicators helping each other's replication in a cyclical topology: I_1 supports the copying of I_2, which helps the replication of I_3, etc; the last member of

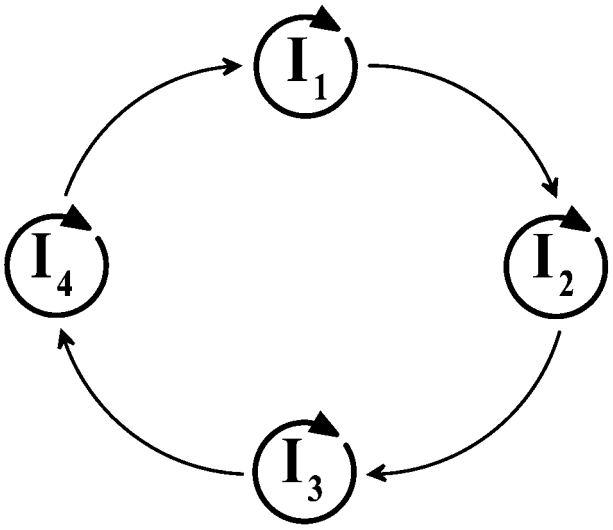


Figure 1. Hypercyclic coupling of autocatalytic replicators $I_1,\ldots,I_4$.

the cycle assists the replication of I_1 (Eigen and Schuster, 1979) (Fig. 1). The chemical nature of the help given to the next member of the hypercycle can be direct catalysis. Each member carries two genes: one for the replication, the other for the catalytic help. There is coexistence in the non-spatial model of hypercycle: all the replicators persist in the system and, given an unlimited supply of monomers, the total concentration of the macromolecules admits hyperbolic growth. The problem with this model is its vulnerability to the invasion of two kinds of parasitic mutants: selfish and shortcut-parasites (Maynard Smith, 1979; Maynard Smith and Szathmáry, 1995). A selfish parasite receives specific help from the previous member of the hypercycle, but it does not help the next one. As a consequence, the mutant sucks off all resources from the cycle, and ultimately only the parasite remains. Shortcut parasites receive help like any other member of the hypercycle, but they support the replication of a farther member instead of the next one, thus reducing the number of coexistent replicators. A series of shortcut parasites, again, reduces the size of the hypercycle to a single autocatalytic member; most of the information carried by the system as a whole is lost (Fig. 2).

There are two known ways of keeping parasitic mutants off the hypercycle: 1) by wrapping it in a membrane vesicle and letting vesicle-level selection eliminate those 'micro-spheres' carrying deleterious mutants (Eigen *et al.*, 1981; Maynard Smith and Szathmáry, 1995), or 2) by having the hypercycle work on a two-dimensional surface, with the assumption of limited replicator mobility and local interactions among the macromolecules (Boerlijst and Hogeweg, 1991; Cronhjort and Blomberg, 1995). The wrapped hypercycle solution would necessitate a still missing, feasible explanation for the appearance of membrane units. Hence the surface dynamics of hypercycles seems more realistic. Here we assume (Boerlijst

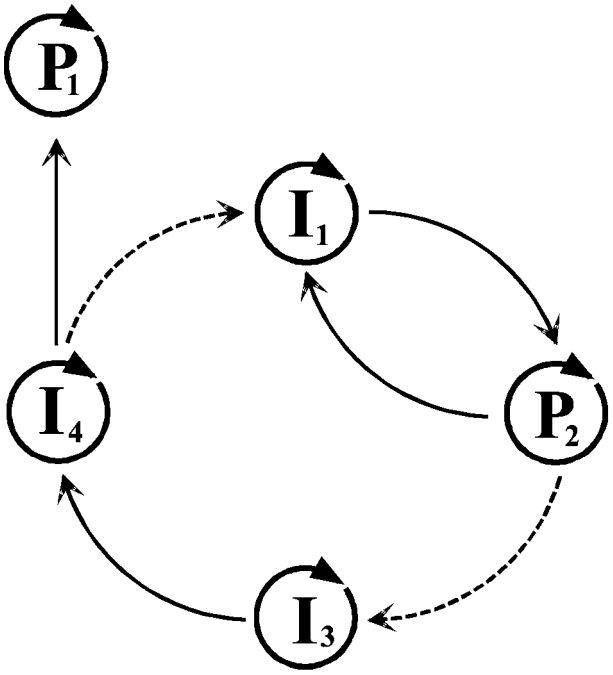


Figure 2. Parasites of the hypercycle. P_1: selfish parasite; P_2: shortcut parasite.

and Hogeweg, 1991) that the macromolecules can help each other's replication if they happen to be neighbours on a square lattice representing a mineral surface. Diffusion has a limited potential of rearranging the replicator pattern: even rather strong diffusion cannot confuse the conspicuous rotating spiral waves emerging as spatial manifestations of cyclic mutualism. The most interesting feature of the spatial hypercycle is its resistance to selfish parasites: the rotating waves drive the parasite out of the spiral, at the margin of which the parasite gets lost. The same conclusion has been drawn with only quantitative differences using a reaction-diffusion model of the hypercycle (Cronhjort and Blomberg, 1995), but the effect of shortcut parasites has not been studied in spatial models yet: it is expected that shorter hypercycles select out the longer ones, leading to only four members remaining, hence the spiral waves and the resistance against parasites would disappear.

2.2. METABOLIC COOPERATION ON A SURFACE

Another class of ecological solutions has been built on the trait group model of Wilson (1980), assuming that replication requires the obligatory metabolic cooperation of the replicators, each of which contributes to the production of the monomers they are all built of. The replicators act as metabolic enzymes catalysing essential reactions in the metabolic network, thus the lack of any one of

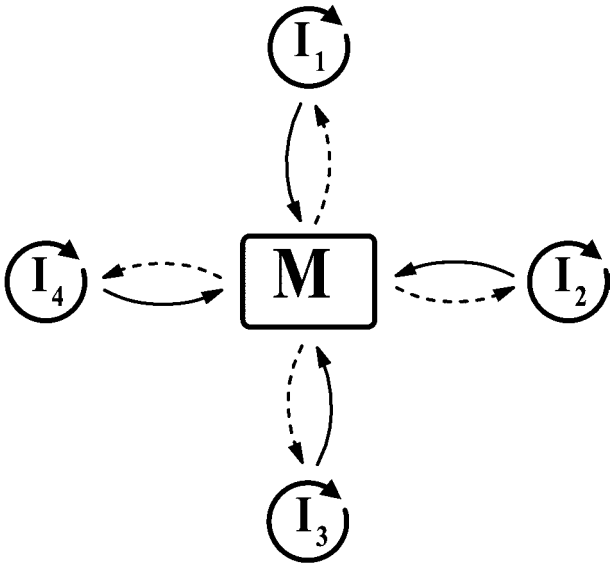


Figure 3. Metabolic coupling of autocatalytic replicators. Replicators $I_1, \ldots, I_4$ contribute to the common metabolism M, and metabolism produces monomers for the replication.

them immediately stops the monomer supply and thus also replication (Szathmáry and Demeter, 1987; Czárán and Szathmáry, 2000). In the metabolic model each replicator aspecifically contributes to the replication of all macromolecules in the system, including itself (Fig. 3).

This is a striking difference compared to the hypercycle in which each member specifically helps the replication of just another one (Fig. 1). Since metabolic help is aspecific, the fastest (of highest replication constant k_i) macromolecule excludes all other replicators from the system, and, metabolism thus having broken down, eventually all of them die out in a non-structured model. This is easy to see from the non-spatial (mean-field) representation of the metabolic system:

$$\frac{dx_i}{dt} = x_i[k_i M(\mathbf{x}) - \phi(\mathbf{x})], \tag{1}$$

where x_i is the concentration and k_i is the replication rate of replicator type i,

$$M(\mathbf{x}) = \left[\prod_{i=1}^{n} x_i\right]^{1/n}, \tag{2}$$

is the metabolic function of $\mathbf{x}$, the vector of replicator concentrations, and $\phi(\mathbf{x})$ is an outflow function keeping the total concentration of the replicators at a constant value. The metabolic function has its maximum at the most even distribution of $\mathbf{x}$. $M(\mathbf{x})$ is changing in time as $\mathbf{x}$ changes, but it is the same for each replicator at any given instance, therefore the fastest replicator never gets slower than the others

– hence the competitive exclusion and the consequent collapse of the mean-field model.

Spatial structure introduced as replicator compartmentalisation into dividing membrane vesicles saves the metabolic system through the *stochastic correction principle* (Szathmáry and Demeter, 1987; Grey *et al.*, 1995; Zintzaras *et al.*, 2002). If each vesicle contains a finite (small) number of macromolecules which are then divided at random into two daughter vesicles through simple vesicle fission, then stochastic drift will re-create metabolically optimal or nearly optimal replicator distributions in some of the vesicles. These will be selected for, thus their offspring will shift the overall replicator distribution towards the metabolic optimum in the next generation. The stochastic corrector mechanism can maintain a viable equilibrium of the replicator types, and, given the metabolic system in the vesicles, it is easier to imagine where the membrane units might come from than it is in the hypercycle model.

Two-dimensional spatial structures have also been shown to keep the metabolic system coexistent (Czárán and Szathmáry, 2000). The 'prebiotic pizza' version of the model is a cellular automaton on an arena of a square lattice. We assume that the macromolecules are reversibly bound to a mineral surface. Each cell of the automaton represents a binding site. The criterion of replication for a bound macromolecule is to be complemented by at least one copy of each of the other replicator types within a certain distance called the *metabolic neighbourhood* (Fig. 4a), so that monomers be supplied at sufficient concentrations locally. Any replicator that has a complete metabolic neighbourhood around can attempt replication into an adjacent empty site. The attempts from within the replication neighbourhood (Fig. 4b) of an empty site are weighted by the product of the replication constant of the corresponding macromolecules and the local value of the metabolic function.

Specifically, the claim C_s of replicator s within the replication neighbourhood of an empty site is

$$C_s = k_s \left[\prod_{i=1}^{n} f_s(i) \right]^{1/n}, \tag{3}$$

where $f_s(i)$ is the copy number of replicator type i within the metabolic neighbourhood of s. The probability that s will actually occupy the focal empty site of the replication neighbourhood is

$$P_s = \frac{C_s}{C_e + \sum_{\ell} C_\ell}, \tag{4}$$

in which C_e is the (constant) claim of the empty site to remain empty. Here ℓ runs through the replication neighbourhood of the empty site.

The cellular automaton is updated randomly; each hit of an occupied site results in the elimination of the resident macromolecule with a probability d, or in no

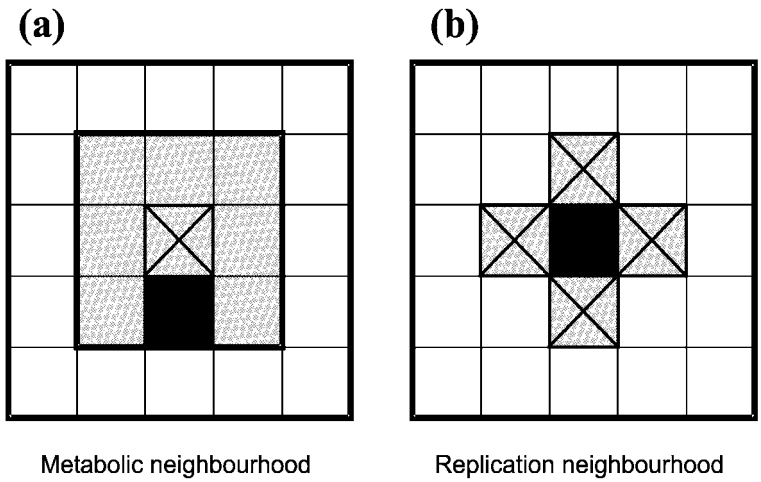


Figure 4. The metabolic (a) and the replication (b) neighbourhoods. Black sites are empty, grey sites constitute the neighbourhood, X labels potential 'mothers'.

change with probability $1 - d$. If updating hits an empty site, it can be occupied by a copy of one of its 4 orthogonal neighbours according to the probability scheme described above.

Each replication step is followed by δ diffusion steps according to the algorithm of Toffoli and Margolus (1987). In other words, δ scales the speed of macromolecule diffusion on the lattice.

Some typical simulation results are shown in Fig. 5. The conclusions of the simulations are the following:

1. the spatially explicit metabolic model warrants the coexistence of the replicators in a wide range of its parameter space;
2. given a fixed number of coexistent, metabolically cooperating replicators (system size), increasing the size of the metabolic neighbourhood is either deleterious or it is beneficial to replicator coexistence;
3. likewise, given a fixed size of the metabolic neighbourhood, increasing the size of the system is either deleterious or it is beneficial for its persistence;
4. more intensive mixing (faster diffusion) is always good for coexistence.

The possibility of coexistence in the spatial model may seem surprising at first glance, knowing that the mean-field version of the model is never persistent. The explanation can be based on a peculiar type of spatial group selection principle. Replicators with small replication constants k_i are rare, and they take a metabolic *advantage of their rarity*. The metabolic neighbourhoods of the rare type replicators are often complete, because they only need the presence of the more common types nearby for that. Thus the rare type has a disproportionally increased likelihood of being replicated. More common types have little chance to be complemented by the rare type, so they have worse prospects of replication. This

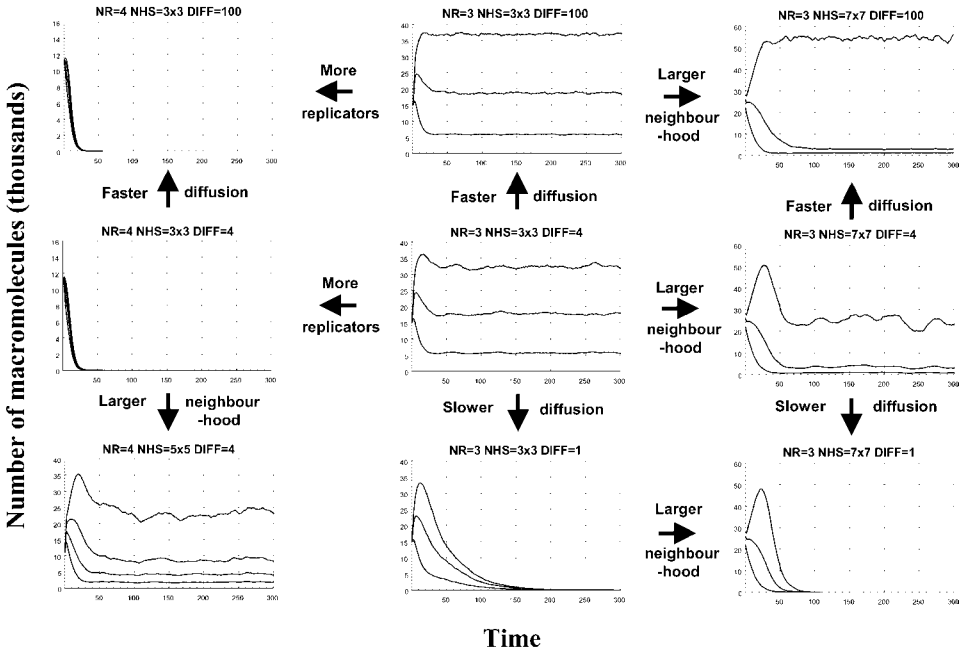


Figure 5. Typical outcomes of the simulations of discrete metabolic networks. Labelled arrows between the panels indicate changes in the parameter values. NR: number of replicators (i.e., system size); NHS: metabolic neighbourhood size (i.e., length of the side of the metabolic neighbourhood); DIFF: number of diffusion steps per generation.

means that the relative chance of replication increases for the rare types and decreases for the common ones, thereby promoting the persistence of the system as a whole. The situation is similar to that of the wrapped metabolic model: members of the metabolically efficient neighbourhoods replicate more, therefore they become over-represented in the overall replicator population due to group selection.

The disadvantage of commonness decreases with larger metabolic neighbourhoods, because larger neighbourhoods are more likely to contain a metabolically complete replicator set including the rare types, so increasing metabolic neighbourhoods are disadvantageous for a persistent system. On the other hand, if the system was too big to fit into a metabolic neighbourhood, then increasing the size of the latter promotes coexistence.

The positive effect of diffusion seems to contradict the expectation suggested by the mean-field model that assumes complete mixing. However, it is not perfect mixing that kills the mean-field system, but the assumption of global interactions: all macromolecules interact with all others, so the competitive advantage of the faster replicators does not vanish until the last copy of the rare type disappears. At that point, the mean-field system collapses. It is in fact the more realistic assumption of local interactions among discrete macromolecule objects that keeps the spatial model coexistent. In this sense, the mean-field system is approached

in the limit by increasing the size of the metabolic neighbourhood, not by more intensive diffusion. Diffusion helps the rare types disperse far apart in the lattice, thus avoiding competition among themselves for empty sites - hence its positive effect on coexistence. If the density of replicators is low within the lattice, diffusion quickly kills off the system, however. This can be regarded as a spatial manifestation of the Allee-effect (Allee, 1938), that is, the decreased reproduction capability of sexual species at low population density. We have summarized these results in Table I.

TABLE I

Summary of the effects of parameter changes on persistence

Parameter change	Direct effect	Effect on persistence
NR ↑	↓ Chance that metabolic neighbourhood contains a complete set of replicators	↓
NHS ↑	↑ Chance that a metabolic neighbourhood contains a complete set of replicators	↑
	↓ Advantage of rarity	↓
DIFF ↑	↑ Spatial mixing (high replicator density)	↑ ⋆
	↑ Spatial mixing (low replicator density)	↓ ⋆⋆

⋆ The system approaches a trait group mechanism (group selection on the neighbourhood configurations) (Wilson, 1980).

⋆⋆ Replicators are dispersed apart (quickly exterminates hopeless systems).

Given the possibility of persistence for the spatial metabolic system, the question of its resistance to possible parasitic mutants naturally arises. The only conceivable parasite of the metabolic system is a replicator that consumes the monomers produced by cooperating members, but itself does not contribute to metabolism at all (Fig. 6).

Introducing such a mutant does not harm system persistence even if the replication rate k_p of the metabolic parasite is much larger than that of the fastest cooperator, but it coexists with the cooperators at a density well below than expected from its replication constant (Fig. 7).

The reason for the modest behaviour of the parasite is that it needs the presence of all the cooperators within its metabolic neighbourhood for replication, but the cooperating members do not need the parasite around for that. That is, the parasite suffers all the disadvantages of rarity and of commonness, without actually enjoying any of the advantages that cooperators can rely on.

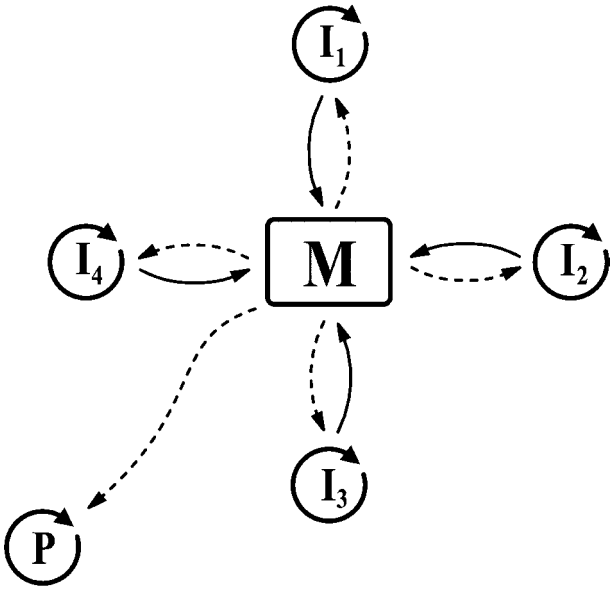


Figure 6. The parasite of the metabolic replicator system: it uses the monomers supplied by metabolism, to which it does not contribute at all.

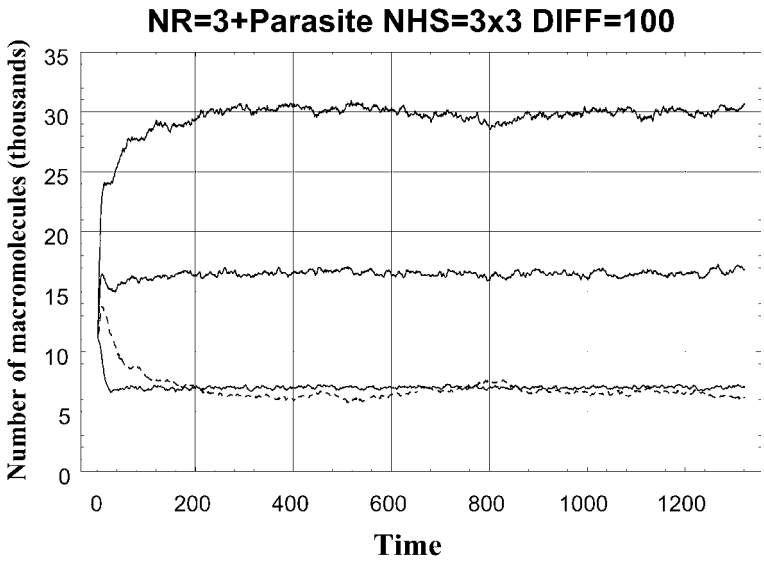


Figure 7. Coexistence of cooperators and the metabolic parasite. Replication constants: $k_1 = 2, k_2 = 4, k_3 = 6, k_p = 8$. Solid lines are cooperating members, dashed line shows the parasite's performance.

Since parasites can be persistent in the system indefinitely, they might have ample possibilities to mutate. Such mutant parasites might have played a central role in prebiotic evolution as suggested in the Discussion.

2.3. HYPERCYCLE AND METABOLIC COOPERATION IN A HETEROGENEOUS SURFACE

Real mineral surfaces are not homogeneous, so it is of fundamental importance to study whether the hypercycle and metabolic models are viable in spatially heterogeneous environments, that is, whether coexistence and resistance against parasites remain valid.

First let us investigate the spatial hypercycle model from this point of view. As we have seen earlier, the coexistence of replicators manifests itself as spiral structures consisting of waves of subsequent members of the cycle. This peculiar spatial structure enables the hypercycle to resist invasion by selfish parasites.

The two important parameters of this model, determining the scale of the emergent spatial patterns, are diffusion and death rates. Spatial heterogeneity of any of these parameters is expected to influence the persistence of the spiral waves required to drive back the parasites. We modified the original model by including spatially heterogeneous decay rate. Spatial heterogeneity appears as a patchy distribution of sites with high and low decay rates. Higher decay rate can be interpreted as a higher desorption rate from the surface. The hypercycle persists on such a heterogeneous surface as well, but parasite resistance is *lost*. Due to the varying width of the spiral arms, inoculated parasites are able to travel between waves consisting of molecules providing catalytic support separated by waves of unsupporting types (Fig. 8). Consequently, some time after an invasion event only the parasites remain in the system (Fig. 9). In a homogeneous environment made of only sites with low or high decay rates the hypercycle resisted parasites.

In the surface metabolism model there is no emergent mesoscopic pattern like the spiral waves, since a relatively homogeneous spatial distribution of the replicator types is necessary for many neighbourhoods to contain a metabolically sufficient set of macromolecules. Consequently, there is no sound reason to expect that spatial heterogeneity of any of the parameters would destroy coexistence or deprive parasite resistance. This view is supported also by computer simulations, we observed qualitatively the same results for the heterogeneous system as in the homogeneous one (see Fig. 5).

2.4. PREBIOTIC SOUP: OPEN CHAOTIC FLOW

The usual model of chemical or biological activity in fluids implies a well-mixed environment, where the chemical constituents or distribution of biological entities can be modeled with their averaged concentration (e.g. Eigen and Schuster, 1979). Recent studies (Aref *et al.*, 1989; Rom-Kedar *et al.*, 1990; Shariff *et al.*, 1991; Jung *et al.*, 1993; Ziemniak *et al.*, 1994; Péntek *et al.*, 1995a; Péntek *et al.*,

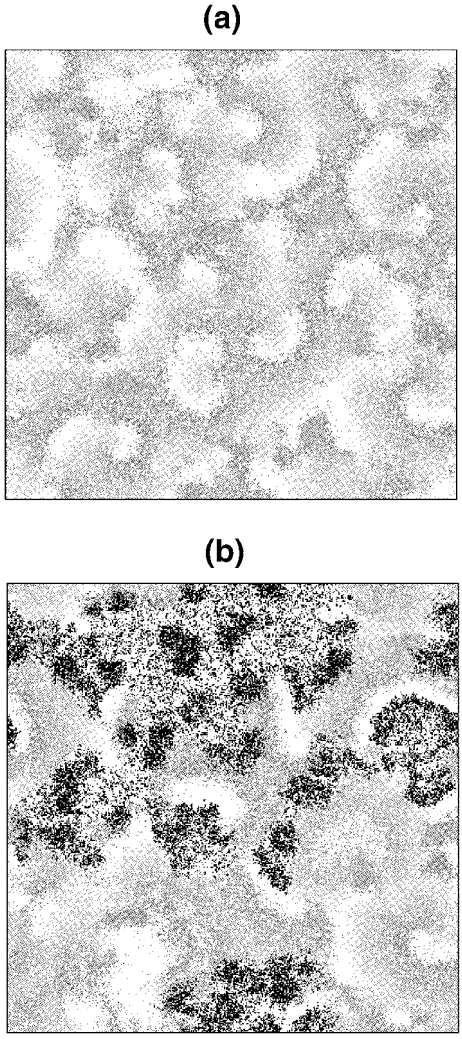


Figure 8. Invasion of the spatial hypercycle by parasites in a heterogeneous environment. The replication surface consisted of 400x400 sites with high and low death probability as 100 blocks of size 20x20 with P_d = 0.05 scattered in the see of unfavorable sites with P_d = 0.2. All the other parameters are the same as those used by Boerlijst and Hogeweg (1991). (**a**) The spatial pattern of a nine membered hypercycle without parasites after 800 steps. Members of the cycle are coloured different shades of gray. (**b**) The system was inoculated by 1000 parasites in the upper left quarter of the lattice after 1000 steps. Invasion of parasites (black) 1200 steps after inoculation is shown. Later the hypercycle completely annihilated. In a homogenous environment consisting of only sites with P_d = 0.05 or P_d = 0.2, respectively, the hypercycle drove back the parasites which remained present only as small cysts.

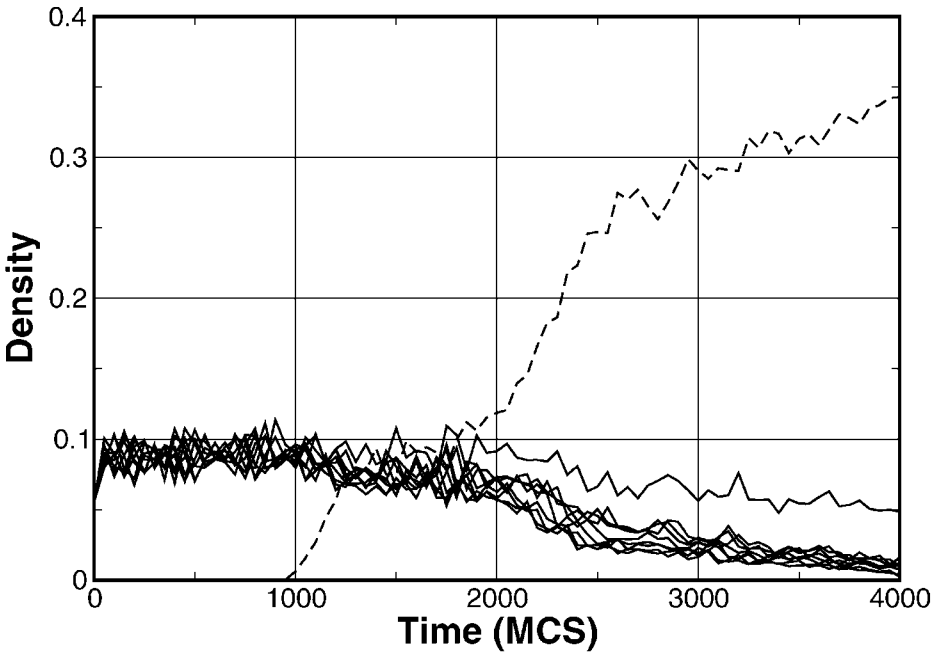


Figure 9. Time evolution of hypercyle in the inhomogeneous environment. Selfish parasites are injected into the system at the 1000th Monte Carlo cycle. Its density increases rapidly (scattered line) while densities of cooperators decrease continously (solid line). The system collapses after some thousand Monte Carlo cycles.

1995b; Sommerer *et al.*, 1996; Toroczkai *et al.*, 1997; Károlyi and Tél, 1997; Rom-Kedar and Poje, 1999) in the field of *open* hydrodynamical flows shed light on systems where *imperfect mixing* is an inherent property of the advected substances. Flows are considered open when there is a certain region of observation where the hydrodynamical flow can transport the advected entities in and out, but the recirculation time between two subsequent entries is significantly longer than the lifetime of the advected entities. Examples for open flows are rivers, certain limited regions of the ocean, deep sea hot springs, etc. When the flow in the region of observation depends on time, even if this time-dependence is simple (non-turbulent), the advected entities typically show very complicated, *chaotic* motion (Ottino, 1989; Péntek *et al.*, 1995a; Károlyi and Tél, 1997). As the open flow will sooner or later wash them out of the region of observation, this chaotic behaviour is necessarily *transient* (Tél, 1990), as opposed to the *permanent chaos* observed in flows restricted to *closed* containers (Aref, 1984; Chaiken *et al.*, 1986; Ottino, 1989; Sommerer and Ott, 1993; Giona *et al.*, 1999). Such transient chaotic behaviour, leading e.g. to patchy distribution of phytoplankton, can be observed in the wake of bridge pillars in rivers, in strong sea currents downstream islands (Arístegui *et al.*, 1997; Barton

et al., 1998), and also in laboratory experiments (Sommerer *et al.*, 1996). Transient chaos is always associated with strong, but imperfect mixing; in flows it means that the advected entities *cannot* be satisfactorily modeled by their averaged concentrations.

For simplicity, we deal with two-dimensional flows only, but the results can be generalized for three dimensional flows. When an ensemble of advected entities reaches the observation region, it typically gets 'trapped' there for a long time while following chaotic orbits, then they trace out a complicated filamentous *fractal* pattern (Péntek *et al.*, 1995a; Péntek *et al.*, 1995b; Sommerer *et al.*, 1996; Toroczkai *et al.*, 1997; Károlyi and Tél, 1997) (see Fig. 10). Mixing is very efficient close to the fractal filaments, but weak elsewhere, resulting in an interesting pattern (Toroczkai *et al.*, 1997) where the chemical or biological activity takes place (c.f. Fig. 11). (In three dimensional flows the fractals form sheets instead of filaments, but the essential findings remain unaltered.) The fact that the activity takes place on a fractal structure, and the advected entities spend an anomalously long time there, modifies the chemical reaction (Toroczkai *et al.*, 1998; Károlyi *et al.*, 1999; Péntek *et al.*, 1999; Tél *et al.*, 2000; Toroczkai *et al.*, 2001) and, correspondingly, the population dynamical equations (Scheuring *et al.*, 2003).

The simplest possible model of replication is the *autocatalytic* reproduction, when the competing entities, in the presence of the necessary constituents, make copies of themselves. In fact, this was probably the most ancient form of reproduction (Maynard Smith and Szathmáry, 1995). In case of autocatalytic reproduction in open flows, fattened-up fractal filaments will be covered by the self-replicating macromolecules. There they reproduce themselves, which leads to a further fattening up of the fractal filaments. This effect is in competition with the outflow from the region of observation. Thus the width ε of the filaments occupied by the macromolecules changes in time according to the equation $\dot{\varepsilon} = cv_R - \lambda\varepsilon$, where c is a geometrical constant, v_R is the velocity by which the width would increase due to self-replication, and λ is the exponential rate of narrowing due to outflow and convergence to fractal filaments. In dynamical systems jargon, λ is the average Lyapunov exponent (Tél *et al.*, 2000). The competition between the reproduction and the outflow typically results in balance, that is, there is an equilibrium filament width traced out by the macromolecules: $\varepsilon^* = cv_R/\lambda$. The fact that the covered filaments form a fractal implies that the area A covered by the accumulated macromolecules scales non-trivially with the filament width: $A \sim \varepsilon^{2-D}$, where D is the fractal dimension of the filamentous fractal. Accordingly, the rate of change of the covered area can be described by the equation (Toroczkai *et al.*, 1998; Károlyi *et al.*, 1999; Tél *et al.*, 2000)

$$\frac{dA}{dt} = -\kappa A + cv_R A^{-\beta}, \tag{5}$$

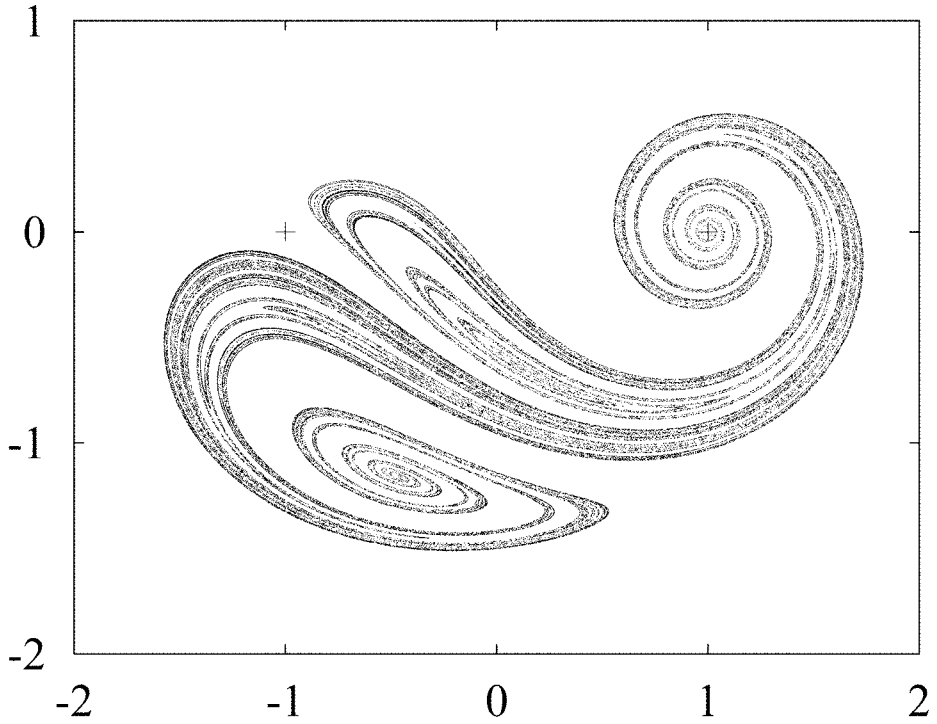


Figure 10. Fractal pattern traced out by passively advected entities, which are trapped in the time-dependent region of the *blinking vortex-sink flow*. The flow consists of two sinking vortex apart from each other and functioning in an alternating manner: during the first half time period of the flow, a sinking vortex drains the fluid at ($x = -1$, $y = 0$), then, during the next half time period, another one works at ($x = 1$, $y = 0$) (Aref *et al*., 1989; Károlyi and Tél, 1997). The black dots indicate the position of the 'trapped' tracers at the time instant when the left sinking vortex closes and the right sinking vortex is about to open. Note the filamentous fractal structure of the black dots, this pattern changes shape periodically in time, it is synchronized to the flow.

where $\kappa = (2 - D)\lambda$ is the *escape rate* describing the exponential rate of outflow along the filaments, and c is, again, a geometrical constant, while

$$\beta = \frac{D-1}{2-D} > 0 \tag{6}$$

is a non-trivial exponent. Exponent β, which depends uniquely on the fractal dimension, characterizes the enhancement of reproduction due to the fractality produced by the advection in open flows. This novel type of 'reaction equation' describing the effects of chaotic advection on autocatalytic processes can be generalized for more complicated reactions or biological activities, with the non-trivial scaling exponent β remaining intact (Károlyi *et al.*, 1999; Scheuring *et al.*, 2003).

One therefore observes that even a simple growth process, such as the autocatalytic reaction subjected to imperfect mixing, follows a nonlinear reaction kinetics.

Figure 11. SEAWIFS image of a phytoplankton bloom at Shetland Islands, May 12, 2000, from the NASA archive. Plankton individuals (light grey) move along a fractal set.

The reason for this nonlinear behaviour can intuitively be explained as follows. On one hand, chemical processes are nonlinear processes with *spatial* character: the reacting molecules have to come in close contact with each other in order for the reactions to take place, which obviously is a spatial constraint. On the other hand the position of the molecules is changing in a nonlinear fashion due to the imperfect mixing dynamics dominating the flow itself. The two spatial nonlinearities thus couple on a *local* level, giving rise to a strong nonlinearity of the reaction kinetics. It is important to repeat the fact that *imperfect mixing* is a necessary condition for

the existence of nonlinear kinetics: for homogeneously mixed components, or pure surface reactions (no mixing at all) the exponent β in Eq. (5) becomes a trivial number, namely, zero for surface reactions ($D = 1$) and infinity for perfect mixing $D = 2$. In both cases the reaction kinetics becomes a linear equation.

An important feature of the nonlinear equation (5) is that it involves the *advantage of rarity* principle: since $-\beta$ is a *negative* exponent, competitors with less abundance (small overall area) will cover the fractal filaments at an enhanced resolution, leading to an increased activity. In other words, covering a fractal with increased resolution implies that smaller amount of material will contribute with an increased second term in Eq. (5). This then competes with the first term, i.e., the decay term (the decay term may contain, besides the loss due to outflow from the mixing region, the spontaneous decay of the species, which only leads to a renormalization of the decay rate κ, leaving the equation unchanged). It is easy to see that this competition will lead to a steady state situation where the decay balances the growth term.

In population dynamics terms it is more natural to talk about the number of individuals, instead of the spatial area occupied by them. Equation (5) can easily be translated into this language by observing that $N = \epsilon_0^{-2} A$, where ϵ_0 denotes the linear spatial extent of a single individual. Then Eq. (5) becomes:

$$\frac{dN}{dt} = -\kappa N + q v_R N^{-\beta} \tag{7}$$

where q is just a factor given by $q = c\epsilon_0^{-2(\beta+1)}$.

The lifetime of the self-replicating macromolecules in the prebiotic ocean was certainly much smaller than the recirculating time to, e.g., a deep-sea hot spring. Hence the prebiotic soup can be assumed to behave as an open flow. This means that the classical studies attempting to describe the prebiotic competitors with their averaged concentrations have to be replaced by the new theory including the effects of the fractality, i.e., the nontrivial spatial distribution of the competitors. This results in the singular scaling exponent β appearing in the equations describing reproduction. Note that when there is only a small amount of self-replicating macromolecules present in the region of observation, the production term in Eq. (5) will largely increase due to the positivity of β. Due to the discrete number of individuals, however, N cannot be arbitrarily close to zero, thus (7) is valid only for positive N. This is the result of the fractal being covered with very fine filaments providing enhanced access to resources, and this might effectively prevent the macromolecules from extinction. This is examined in more details in the next subsection, where the effects of competition are discussed.

2.5. PURE COMPETITORS IN OPEN CHAOTIC FLOW

The previous section studied the case of a single replicator type occupying a niche amidst imperfectly mixing flow conditions. Since we are interested in information integration, our basic questions are whether two or more types of replicators

competing for the same limiting resource in an imperfectly mixing chaotic flow can coexist, and whether this coexistence is robust. In the following we answer these questions, presenting a spatially explicit model of this phenomenon and using mostly heuristic mathematical arguments. For the more detailed, accurate, and technical analysis see Károlyi *et al.* (2000), and Scheuring *et al.* (2003).

Before discussing the consequences of imperfect mixing on competition, it is worth briefly giving the traditional equations governing autocatalytic processes in a well-mixed environment. In a fixed region of observation they are:

$$\frac{dN_i}{dt} = \gamma_i R N_i - \mu_i N_i, \tag{8}$$

where N_i denotes the number of individuals of replicator I_i, and R is the instantaneous amount of the resource material in the same region. If the dynamics of resource is much faster than the dynamics of the competing replicators, then the former can be considered to be in a quasi-stationary state: $dR/dt = 0$. The equation for resource R is then

$$\frac{dR}{dt} = 0 = J - R \sum_i \gamma_i N_i, \tag{9}$$

and J is the constant inflow of resource R into the region of observation. After analyzing (8) and (9), one can easily see that replicator with lower γ_i/μ_i ratio of replication and death rates would be outcompeted, and thus stable coexistence is impossible.

As in the traditional model we consider a simple kinetic model of replication and competition with passively advected point like replicators of type I_1 and I_2, multiplying themselves instantaneously after certain replication times. This kinetic model is similar in spirit to the one used by Metcalfe and Ottino (1994) for modeling chemical reactions in *closed* flows. In our *open* flow, there is a constant inflow of the resource material A into the system, for which the different replicators I_1 and I_2 compete. In addition, there is a spontaneous mortality of individuals. Therefore two autocatalytic processes $R + I_1 \to 2I_1$, $I_1 \to R$ and $R + I_2 \to 2I_2$, $I_2 \to R$ represent the replication and competition process in our model in an imperfectly mixed environment. The kinetic coefficients γ_i (μ_i) give the birth (mortality) rate as before, i.e., number of individuals being born (dying) in unit time in a homogeneous environment. As we emphasized above, according to the traditional theory, the replicator with lower γ/μ ratio would be outcompeted in a *well mixed*, homogeneous environment, there is no way for the information to be stored in different types of replicators.

An important feature of the advection dynamics is its purely deterministic nature. This implies that we work in the limit of weak diffusion and assume that the mutual diffusion coefficient between any pair of the constituents is small.

The competitive dynamics starts with the full surface occupied by the background material R. Initially, we place two droplets of individuals from replicators

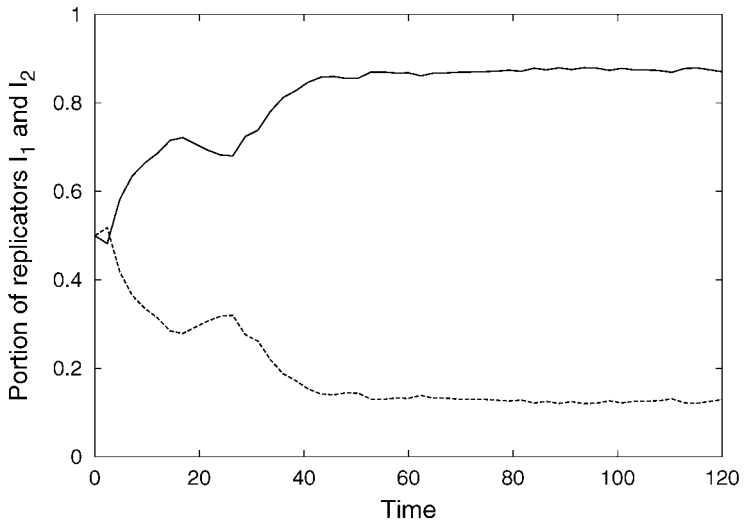


Figure 12. The portion of replicators is shown in the region of observation. The flow considered was a model of the *von Kármán vortex street*. In this case, there is a time-periodic detachment of vortices in the wake of a cylindrical obstacle placed into a uniform flow. This *open*, *time-periodic* (with period T) flow traps the advected competitors in the wake of the obstacle along a fractal curve. Time is measured in units of T. Replicator I_1 make copies at time lags $0.6T$, I_2 at time lags $0.8T$. Both replicators decompose with the same probability 1/10.

I_1 and I_2 into the flow in the inflow region with I_2 being the weaker competitor. We monitor the number of individuals present in the region of observation during the competition process. The iteration consists of two steps. The first step models the advection of the replicators on the chosen grid, while the second one is the instantaneous replication occurring on the same grid. A molecule can replicate only if at least one of its neighbours contains resource R. The faster replicator creates more copies of itself into the neighbouring cells than the slower one within unit time. The advection phase also includes the random death of individuals with a probability specific to its type I_1 or I_2. There is thus a loss of the individuals due to the advection and the finite life-time, but also a gain due to the reproductions.

After an initial rapid increase, the numbers of I_1 and I_2 replicators become synchronized with the flow, and the system settles into a dynamic steady state. We emphasize that replicators coexist in spite of their very different γ/μ ratio, see Fig. 12, even if more than two different type of replicators are present (Károlyi *et al.*, 2000). The magnified picture demonstrates the filamentous structure where competitors accumulate, and also the imperfect mixing in this open chaotic flow, see Fig. 13.

Results of numerical simulations are reinforced by the analytical description based on the observation that replicators I_1 and I_2 will occupy the filaments of a fractal, and along one such filament they become organized along more or less parallel stripes (Scheuring *et al.*, 2003). This results in two types of reactive inter-

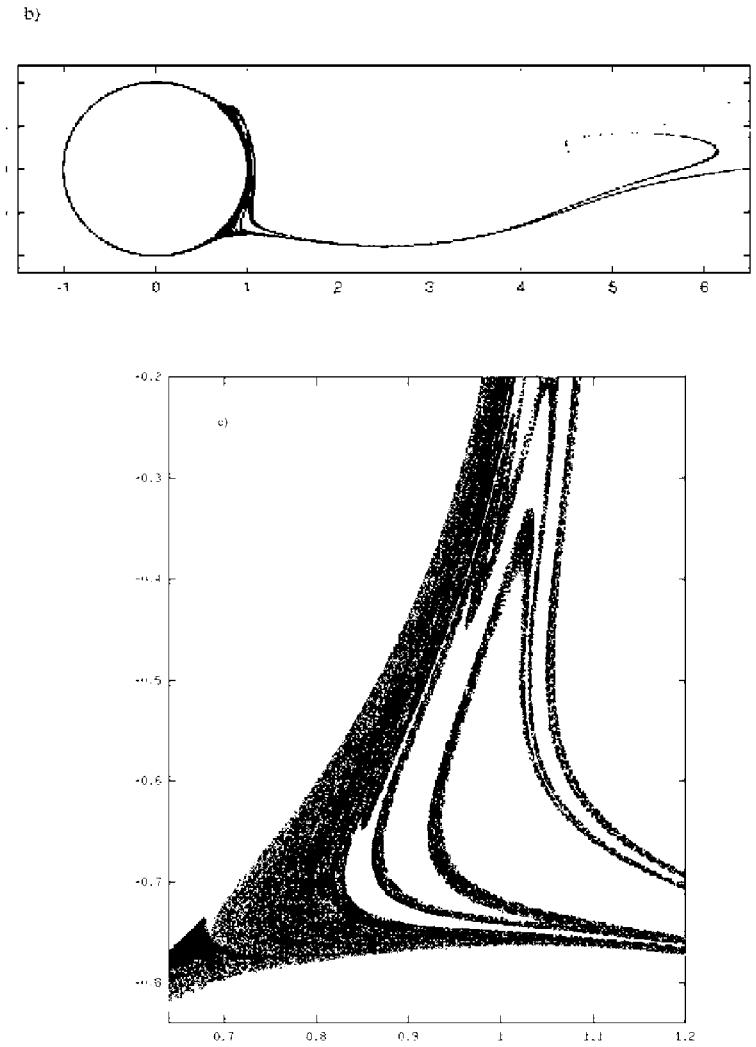


Figure 13. (a) Spatial distribution of replicators I_1 (black) and I_2 (gray) after 24 periods of the flow. The flow is a model of the von Kármán vortex street: the circular region is the cylindrical obstacle, the flow is directed from left to right. The model flow corresponds to a Reynolds number of about 250. The parameters are such that $\gamma_1/\mu_1 = 3/4 = 3\gamma_2/4\mu_2$, and thus I_1 is the weaker species. (b) The magnified rectangular region of (a). Replicators I_1 (black) and I_2 (gray) are distributed along filaments of a fractal.

faces, namely, the R-I_1 and R-I_2 interfaces. To characterize the distribution of each replicator within the filaments of this fractal, the probability distributions p_i are introduced, which express the probability that one finds an active R-I_i interface on the border of a randomly chosen filament. Naturally these probabilities will depend on the average filament width for each replicator and their replication rates. Simple

dimensional reasoning, and symmetry arguments suggest that a family of solutions can be obtained in the form of a generalized power-law:

$$p_1 = \frac{z^\alpha}{z^\alpha + \omega}, \quad p_2 = 1 - p_1, \tag{10}$$

where $z = \varepsilon_1/\varepsilon_2$ is the ratio of the average total stripe widths of the replicators covering the fractal filaments, $\omega > 0$ depends on the replication rates, and $\alpha > 0$ depends on the details of the flow (Scheuring *et al.*, 2003). In the range of $0 < \alpha < 1$ the smaller population is less probable on the boundary but yet with a weight which is weaker than linear in the widths. For $\alpha = 0$ there is no width-dependence at all, the probabilities p_i are constant. The case $\alpha = 1$ and $\omega = 1$ corresponds to a homogeneous mixing within the stripe of width ε. For $\alpha > 1$ a superdominance is described.

It follows that for $0 < \alpha < 1$ *coexistence is stable*, and for $\alpha > 1$ it becomes *unstable*. The detailed examination of the population dynamical equations and of the stability of the coexistence fixed points is to be found in the Appendix, see also Scheuring *et al.* (2003).

The case $\alpha = 1$ is special. For $\alpha = 1$, and $\omega \neq v_1/v_2$, the non-coexistence point is the only fixed point, and it is stable. Having $\alpha = 1$ with $\omega = v_1/v_2$ implies that there is an infinity of fixed points with marginal stability. Thus, stable coexistence is found in the

$$0 < \alpha < 1 \tag{11}$$

regime. These results were found to be in complete agreement with the numerical calculations presented by Scheuring *et al.* (2003), namely that a measurement of p_1 gave indeed the generalized power law shown in (10), and whenever coexistence was observed numerically, it was indeed consistent with condition (11), and when coexistence was not found, we had $\alpha \geq 1$. Interestingly, $\alpha < 1$ is formally analogous to the parabolic replication (Szathmáry and Gladkih, 1989; Sievert and von Kiedrowski, 1994; Wills *et al.*, 1998), however this sub-exponential law of concentration increase followed from the peculiar mixing structure of the flow not the self-inhibition of replicators.

2.6. Metabolic cooperation in open chaotic flow

In the previous subsection we have shown that coexistence of competing autocatalytic macromolecules is possible in open flows. To promote this finding, it is necessary to check whether coexistence is possible for more complex interactions of advected replicators, e.g., when macromolecules are linked by metabolic cooperation. Intuitively one would expect that including some sort of cooperation among purely competing autocatalytic macromolecules can only increase the chance of coexistence. This is, however, not entirely true; in fact, we have seen that in a well-mixed, homogeneous environment the addition of metabolic cooperation does *not* imply coexistence (see Section 2.2).

When metabolically coupled processes take place in open chaotic flows, the inherent properties of the activity are not changed. The details – and the model – of the activity remain the same as in Section 2.2. The macromolecules are distributed on a grid, each grid-cell occupied by at most one macromolecule. The size of the grid-cell models the finite size of the macromolecules. When one such grid-cell happens to be empty, then its neighbours might have the chance to put an offspring into it, but only if the common metabolism in the local neighbourhood supports reproduction. In other words, the macromolecule has the chance to put an offspring into a neighbouring empty grid-cell if in the metabolic neighbourhood of the macromolecule there is at least one copy of each macromolecule types contributing to the common metabolism. The chance of reproduction is proportional to the geometric mean of the number of all types of macromolecules in the metabolic neighbourhood, and to its type-specific replication constant.

The effect of the flow is modeled by moving the existing macromolecules into new grid-cells. The macromolecule in a certain grid-cell is considered to be located in its middle, and is advected passively as an inertia-less particle into a new location during a certain discrete time-step. Then it is replaced to the center of the grid-cell in which it has arrived. These replacements together with the finite grid-size model also a small diffusion.

Thus one step of the advection-reaction-diffusion process consists of the following sub-steps:

1. advection of the macromolecules into new grid-cells during a short discrete time-step. Then they are replaced into the center of the new grid-cell, modeling the effect of diffusion.
2. spontaneous decay of the macromolecules with a type-specific decay constant. This is not an essential part of the simulation, because the outflow already takes the role of decay, nevertheless, it is included for biological reality. Decay is modeled by eliminating all macromolecules with a small type-specific probability.
3. the content of each grid-cell is updated simultaneously according to the metabolic scheme.

The most important factors that determine the actual outcome are community size (number of competitor types), metabolic neighbourhood size (what is the maximum distance across which the common metabolism can give support to a macromolecule), and the parameters of the flow (Károlyi *et al.*, 2002).

We found coexistence (see Fig. 14) in a wide range of the parameter values. The reason for coexistence is twofold. On one hand, the discreteness of the model inherently does not allow a perfect mixing, unlike in models with point-like macromolecules. On the other hand, open flow provides imperfect mixing leading to empty spaces that can be occupied by offsprings, and provides the necessary mixing to avoid large areas lacking any of the macromolecule types.

It was also found that the introduction of *parasitic competitors* (using the resources provided by the common metabolism, but not contributing to it) does

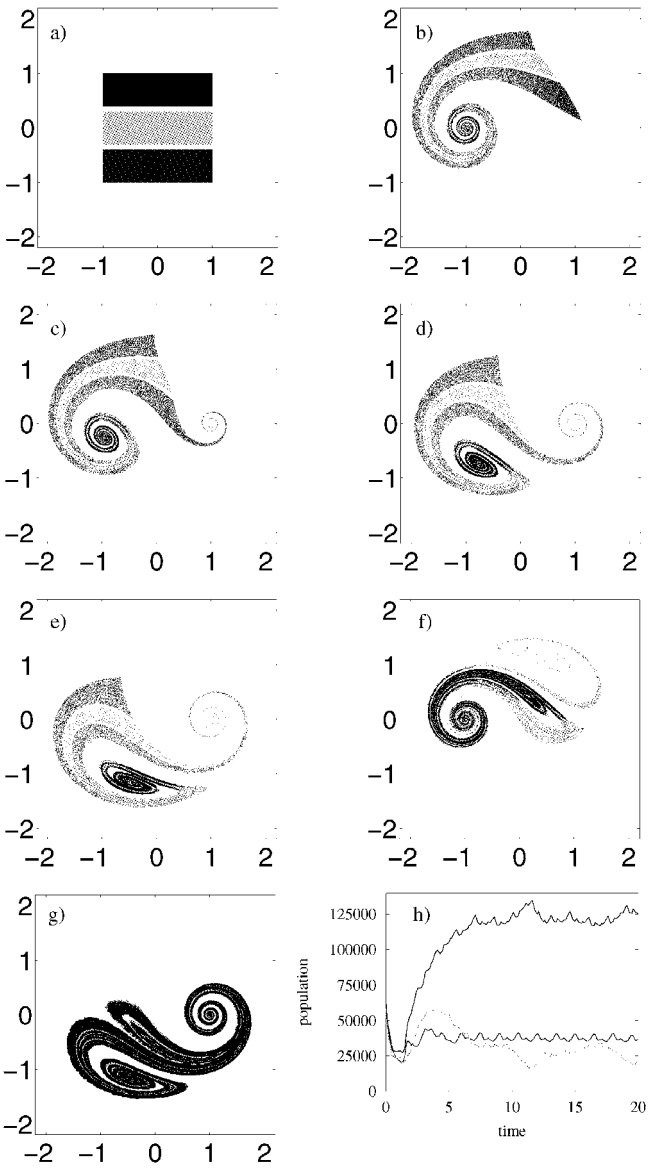


Figure 14. Metabolic network dynamics in the blinking vortex-sink system (the same flow as in Fig. 10). Three different, metabolically coupled competitors are placed into the flow as shown in (a). In the course of time, all of them accumulate on a fractal changing shape time periodically. The snapshots were taken at (a) $t = 0$, (b) $t = 0.5T$, (c) $t = 0.6T$,(d) $t = 0.8T$,(e) $t = T$, (f) $t = 1.5T$, (g) $t = 10T$, while (h) shows how the number of different competitors varies with time (measured in T, the period of the flow). The metabolic neighbourhood size was NHS=10, the replication constants were $k_1 = 3$ (red), $k_2 = 4$ (green), $k_3 = 5$ (blue). Empty sites remained empty with weight $C_e = 2$, decay rate was $\delta = 0.02$.

not toss an otherwise viable system into extinction. This is also explained by the fact that the openness of the flow, through imperfect mixing, provides sufficient amounts of empty spaces for reproduction. Also, when the number of parasitic macromolecules were to increase beyond a certain threshold, this spreading has to slow down as the common metabolism slows down in lack of other, contributing macromolecules. In a well-mixed environment this is connected to the total extinction of the competitors. In case of an imperfectly mixed environment, however, there will always remain localities where the deviation from the average is large enough to serve as a core for recovering.

3. Coevolution of Replicators

Although, as we emphasized earlier, the hypercycle hypothesis is problematic from many different points of view, the catalytic role of ribozymes in replication might have been important in the early evolution of replicases. The basic idea (which is not incorporated in Eigen's original model) is that the longer the RNAs are, the better could be their catalytic activity, fidelity, and template efficiency as well. So, according to the iterative scenario for longer and longer molecules with better and better replicase functions, prebiotic evolution might have escaped the Catch 22 (James and Ellington, 1999; Joyce and Orgel, 1999; Poole *et al.*, 1999). While the possibility of this scenario was analyzed and supported earlier by a simple mathematical model (Scheuring, 2000), the main question remains whether parasites (efficient and fast templates, but ineffective replicases) can ruin the system.

We turn to the the mineral surface again to study the coevolution of a replicator-replicase system. The dynamics on a prebiotic surface are modeled by a cellular automaton with rectangular grid as before. Hypothetical nucleic-acid like macromolecules are built up by different monomers A, B, C, and D. The molecules are adsorbed to the surface, and can interact with each other if they are located on neighbouring sites. They might play one of two roles in an interaction: a molecule is either a template for replication or it acts as a replicase, that is, it helps the copying of the potential template molecule. There is no replication without catalytic aid. The replication process has two important characteristics: speed and fidelity. Naturally, replication speed is proportional to the product of the replicase activity, template efficiency, and monomer supply; and inversely related to template length. For simplicity, it is assumed that template efficiency, replicase activity, and fidelity of molecules are determined by the length of longest blocks of A, B, and C sequences, respectively. It is assumed that all these properties are some increasing sigmoid functions of the length of the corresponding X blocks ($X = A, B, C$) (Fig. 15). The role of monomer D is to mimic neutral mutations, it has no direct effect on replication. The available monomer supply is a linearly decreasing function of total polymer mass on the surface with rate of $1/\mu$. These obviously oversimplified assumptions preserve the most important characteristics

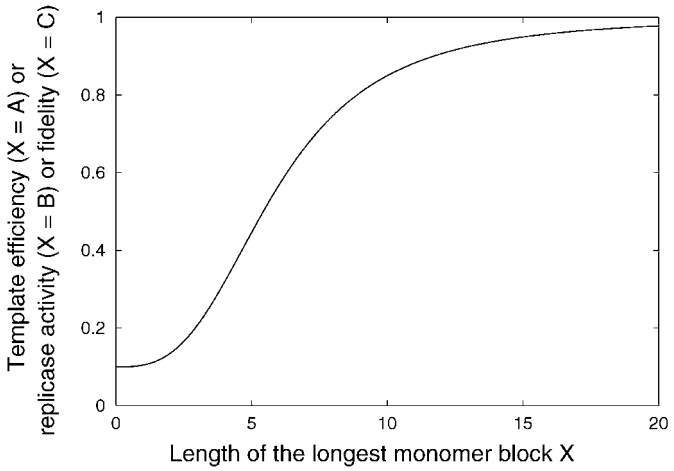


Figure 15. Template efficiency ($X = A$), replicase activity ($X = B$), or fidelity ($X = C$) as a function of the length of the longest monomer block consisting of X. We use the $\alpha_X + (1-\alpha_X) n_X^{\beta_X} / (\gamma_X + n_X^{\beta_X})$ function in the simulations, where n_X is the the length of the longest X sequence ($X = A, B, C$) in the molecule, α_X, β_X and γ_X are positive parameters.

of real macromolecules, namely, that at a given replicator length all these properties of the molecules are in negative trade-offs: a molecule could not be a fast and precise replicase and a good template at the same time. Beside replication, molecules disintegrate spontaneously with a given P_d rate. Replication is not perfect, so there could be point mutations determined by the replicase copying fidelity, and additions or deletions of monomers might occur at a given constant probability (P_{ad}) during replication. Each local site contains only one replicator, so the newborn copies are placed to neighbouring empty sites.

The simulation experiment starts with half of the sites occupied by molecules distributed at random. The replicators were five monomers long with random sequences. The evolution of replicators is followed through many generations. The first observation is that starting from short and diverse branches of oligomers, a complex effective replication machinery emerges (Fig. 16) (Szabó *et al.*, 2002). All the important properties, including replicase activity, fidelity, and template efficiency improved significantly. At the same time the average length of replicators increased as well (Fig. 16). The benefit of increasing template efficiency is obvious. The selective advantage of increased replicase activity and fidelity seems to be surprising. Because of local interactions and low level diffusion present in prebiotic pizza (and in this model, too) an accurate and effective replicase molecule finds similar type of molecules around itself. Thus they form patches of *altruistic* molecules. These altruistic molecules can be parasitised by selfish or even short parasitic templates. But parasites could take advantage from the altruistic replicase molecule only if this selfish type is *locally rare*. Increasing the number of selfish templates in a local aggregated patch decreases their replication success. This is

why parasitic replicators live together with the long altruistic replicators, but they can't destroy the system (Fig. 17). This view is supported by the observation that increasing the diffusion rate decreases the aggregated patterns on the surface and increases the density of parasites. In the mean field counterpart of the model, where each template is copied by the aid of an 'average' replicase, ultimately the whole population dies out (Szabó *et al.*, 2002). Increasing copying accuracy permits the lengthening of replicators, but replication is slower at increased template length. The average equilibrium length is determined by the parameters involved in the sigmoid functions in Fig. 15.

In the light of the results presented here, coevolution of template and replicator function can effectively increase the amount of information potentially carried and can enhance resistance against the template parasites. Replicase evolution and template parasite resistance both require highly constrained mixing, i.e., limited spatial distribution of the replicators.

4. Discussion

According to the theoretical models presented, we can conclude that both the 'prebiotic pizza' and the 'prebiotic soup' could be reasonable habitats for the first replicators. Given the possible mechanisms that might have kept early replicators coexistent, we are now in a position to attempt to construct a feasible scenario of prebiotic evolution starting from the first modular replicators, and leading to the rise of the first proto-cells.

We start with discussing the advantages and the disadvantages of the 'pizza' and the 'soup' concept in general terms. On one hand, an obvious benefit of the 'pizza' scenario is that surface adherence constrains the movement of the macromolecules, which in turn means that the synthesis of long modular molecules implies smaller entropy decrease than it would in fluid medium (Wächtershäuser, 1988). The other benefit of surface reactions is sterical: the macromolecules are fixed and oriented to a certain extent, which enables them being involved in specific reactions that would not be possible in a solution. These effects together make the prebiotic pizza a 'surface catalyst'. Of course, for these benefits to be exploited, surface adherence should be neither too weak nor too strong, which can be true if many different molecules can use the same types of residues for binding. On pyrite surfaces these could be $-PO_3^{2-}$ or $-COO^-$ residues (Wächtershäuser, 1988; Maynard Smith and Szathmáry, 1995).

On the other hand, chaotic flows constrain the molecules on a fractal set where local concentrations can be very high even at low average concentrations. This might have been necessary in the earliest phases of replicator evolution, when the dilute resources of replication were supplied by the chemical environment. Transient chaos typically emerges in any time-dependent open flow, so the physical phenomenon is indeed widespread and robust against external perturbations. Rel-

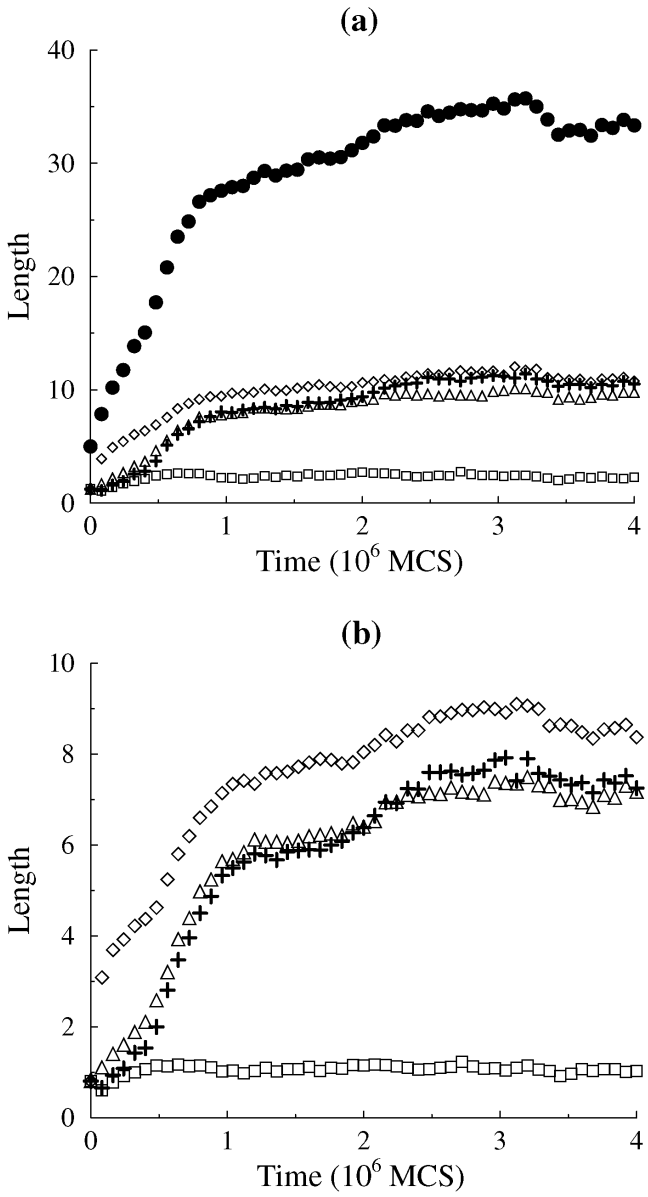


Figure 16. The result of evolution after some millions of Monte Carlo cycles (MCS). (**a**) Time series of average length of replicators (closed circles), the average numbers of monomer A (open diamonds), monomer B (open triangles), monomer C (pluses) and monomer D (open squares). (**b**) Evolution of template (open diamonds) and replicase (open triangles) activities and fidelity of replication (pluses). Parameters: $\alpha_A = 0.1$, $\beta_A = 3$, $\gamma_A = 200$, $\alpha_B = 0.1$, $\beta_B = 3$, $\gamma_B = 200$, $\alpha_C = 0.9$, $\beta_C = 2$, $\gamma_C = 2$, $\mu = 6$, $P_d = 0.001$, $P_{ad} = 0.02$, diffusion is zero.

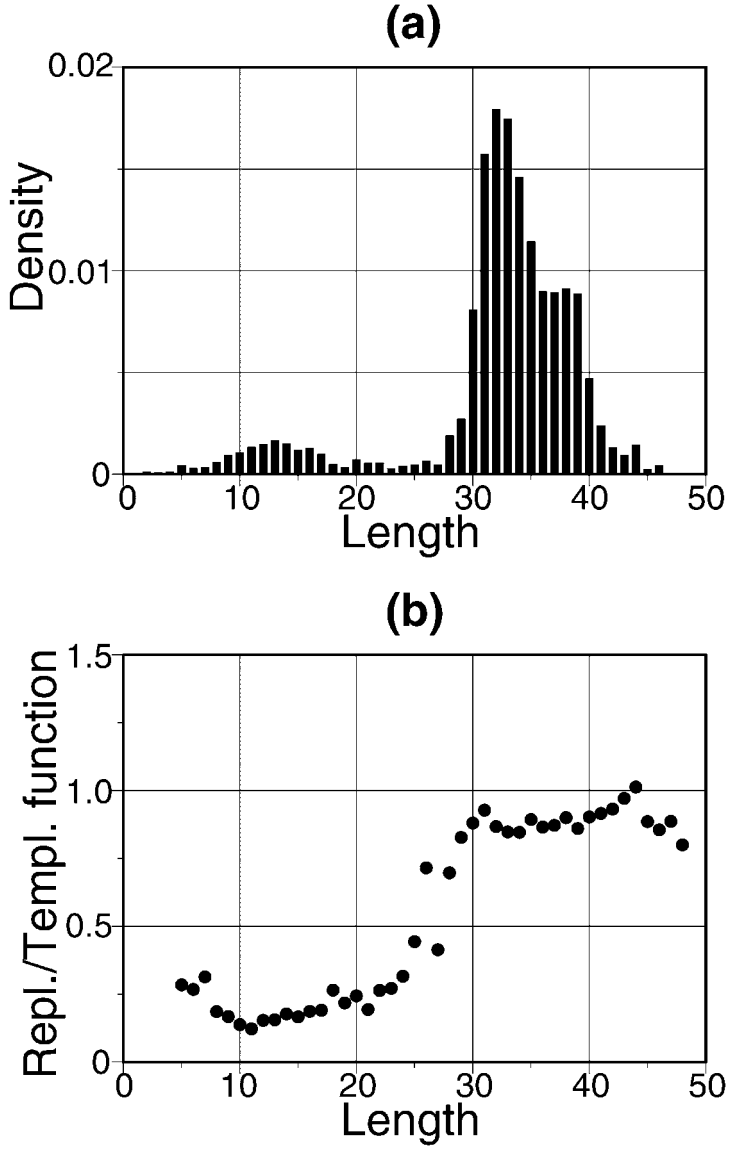


Figure 17. Distribution of replicators at the stationary state. (**a**) Density distribution of replicators in function of length. (**b**) Average ratio of template and replicase function within replicators of different lengths. The smaller peak around length of 13 on figure (**a**) refers to the parasites. Figure (**b**) supports this, showing that these polymers have very weak replicase function while long replicators can act as replicase too. (Parameters are the same as in Fig. 16.)

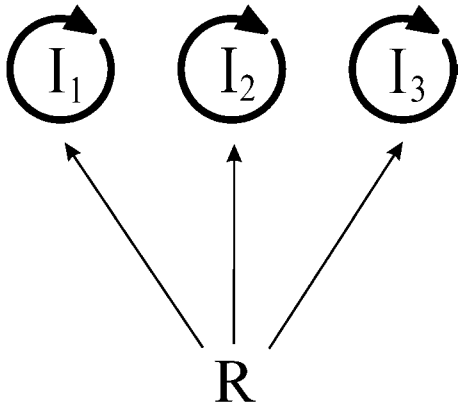


Figure 18. Autocatalytic replicators I_1, I_2, I_3 compete for the same limiting resource R.

evantly, 'black smokers' (deep-sea hydrothermal vents, see Holm (1992)) existed and provided an open flow generating the nontrivial, imperfectly mixed environment.

We suggest that both the 'pizza' and the 'soup' concept might have played a decisive role in prebiotic evolution, either in a parallel or in a sequential manner. One of the possible specific scenarios is the following.

1. Purely competitive autocatalytic replicators used resources supplied by chemical reactions in the environment. The supply of energy and materials for monomer synthesis were external to the replicator system, and they were purely chemical. Chaotic flows may have lead to increased local monomer concentrations, thereby promoting replication on fractal sets. Chaotic flows also enabled the coexistence of otherwise competitively exclusive replicator populations, thus maintaining an initial diversity of the 'genetic information' carried by the primitive community of replicators. At the same time, selection for the better replicators might have taken place either in the turbulent regions of the chaotic flows or on the mineral surface (Fig. 18). We note here that self-association of replicators can lead the coexistence (because of the so-called *parabolic replication*) whenever the replicators are in competitive situation in a well-mixed medium (Szathmáry and Gladkih, 1989; Sievert and von Kiedrowski, 1994). Depending on the kinetic parameters this system can display Darwinian selection as well: either if decomposition is taken into account (Lifson and Lifson, 1999; von Kiedrowski and Szathmáry, 2000; Scheuring and Szathmáry, 2001), or if molecules replicating by enzyme-free autocatalytic ligation (Wills *et al.*, 1998; Stadler *et al.*, 2002). However, the role of spatial processes and constraints for the parabolic replicators have not been studied intensively till yet.
2. The coexisting members of the replicator community were free to mutate. Any mutation improving the cooperation between the replicators was beneficial for

the system, so these cooperative mutants are selected for. There were basically two ways of possible cooperation between replicators: those which helped the production of monomers for replication (metabolic cooperation) and those directly helping replication (replicase cooperation). Ribozymes could catalyze virtually any metabolically important biochemical reaction (Landweber, 1999), but no efficient RNA replicase activity has been found among RNA molecule clones yet. Therefore it seems reasonable to assume that metabolic cooperation might have been more probable to precede the emergence of specific replicase cooperation – metabolism might have been more ancient than specifically catalyzed replication. Our model simulations, not including the effect of the dynamics of small metabolites, suggest that metabolic cooperation might have evolved either in chaotic flows or on a surface: the metabolic replicator system was coexistent in both cases. Metabolic cooperators had two obvious advantages in monomer supply compared to non-cooperating replicators: they were much less dependent on hazardous external factors, and they locally produced monomers directly for their own replication. Since the local manner of monomer production was an important aspect of the cooperation, it might have been more profitable for the cooperators to be fixed to a surface, because it enabled successful 'teams' to prolong their fitness advantage for many generations. The metabolic system could have been indefinitely coexistent with any number of different metabolic parasites which were therefore free to mutate. Deleterious mutations died out, neutral ones persisted, and beneficial ones spreaded. Beneficial mutants could be metabolic cooperators (facultative or obligate) enhancing the performance of the metabolic network, or they could evolve replicase activity. Metabolic cooperators increased system size, and thus the information content of the community of replicators (Fig. 19). (The selection process in this model has remained for future work.)

3. A series of beneficial mutations could lead to a gradual development of replicase activity in one of the metabolic parasites. The new trait thus obtained was good for the (converted) parasite and for the metabolic system alike: it increased the replication rates of all the members of the replicator community. Thus the information content and the efficiency of the system, again, increased in parallel (Fig. 20). Our simulations suggest that this might have occurred on the 'prebiotic pizza', and we conjecture that chaotic flows might have supported such mutants as well, but this latter statement needs theoretical verification.
4. Finally, metabolic parasites could become enzymes that produce membrane units from some intermediate or by-products of metabolism. Once the membrane units were produced, the stage was ready for protocell evolution (Fig. 21). The membrane unit molecules might have formed spherical segments on the surface spontaneously, and the microscopic spheres escaped the surface by abstriction (Wächtershäuser, 1988). Those microspheres containing the complete metabolism-replication-membrane machinery could have started a new way of life in the prebiotic soup, provided that the inflow of nutrients and the outflow

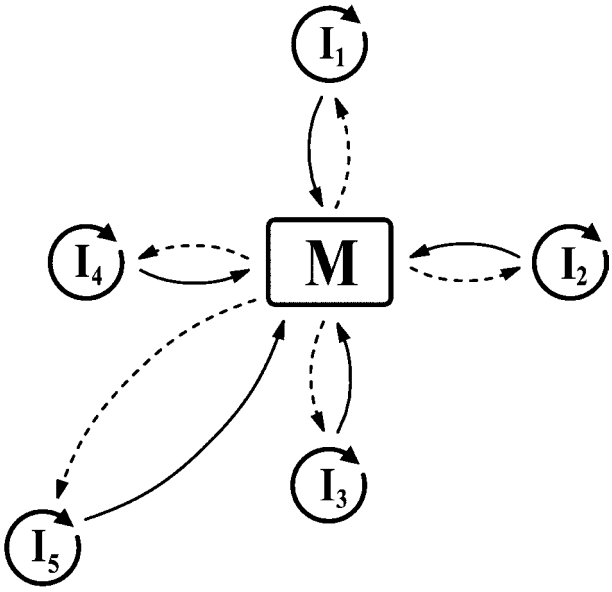
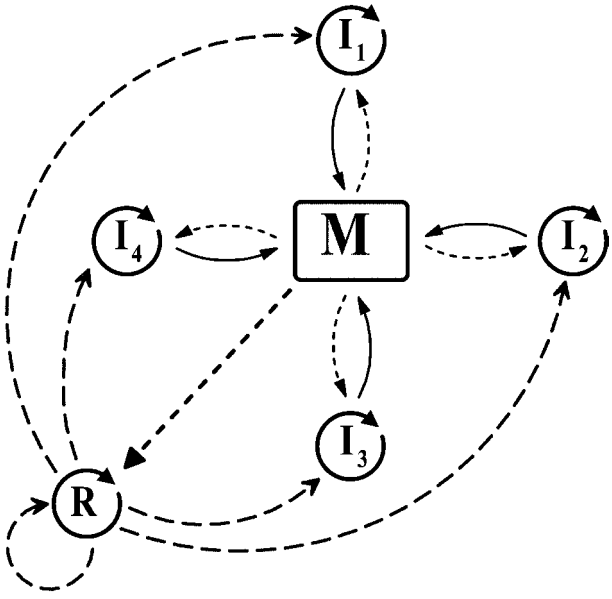


Figure 19. Metabolic parasite converted to metabolic cooperator.



.

Figure 20. Mutant metabolic parasite *R* serves as a replicase (open dashed arrows)

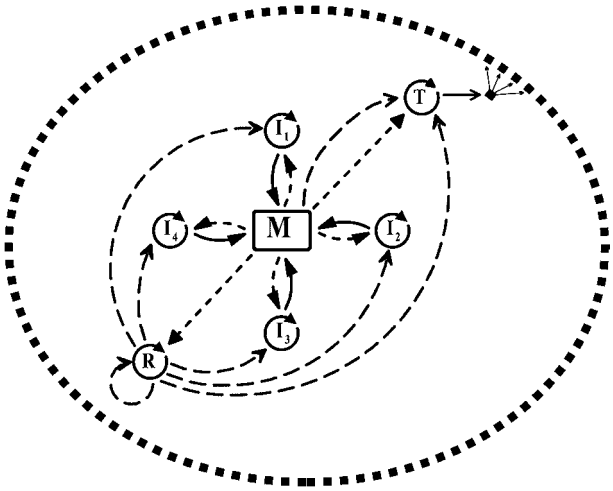


Figure 21. Metabolic system with a replicase and a membrane unit synthesizing ribozyme: the catalytic chemoton.

of waste materials was ensured. This entity was called the *chemoton* by Gánti (1979). Since membrane units were continuously produced by the metabolism, and they built themselves into the membrane surface spontaneously, the chemoton is capable of autonomous division (Gánti, 1979; Koch, 1985).

Eigen's pioneering work on the paradoxical problem of information integration has provoked intensive theoretical work, and it has ultimately led to a sufficiently coherent picture of prebiotic evolution. Our primary suggestion is that the notorious 'pizza' *or* 'soup' controversy could in fact be settled at a 'pizza' *and* 'soup' consensus. Of course the scenario we suggested above is only one of several feasible alternatives, but we believe that surface metabolism and chaotic flow dynamics were key factors in the early stages of replicator evolution.

Acknowledgements

This research was supported by the Hungarian Science Foundation (OTKA) T 037726, T 032423. G.K. and I.S. are indebted to the Bolyai grant for the financial support. G.K. was supported by the Korányi Fellowship of the Thomas Cholnoky Foundation and by the Békésy grant, which is hereby gratefully acknowledged. Z.T. was supported by the Department of Energy, under contract W-7405-ENG-36.

Appendix – Population Dynamics in Open Flows

On the level of stripe dynamics, the equations for the evolution of the partial widths can simply be written as:

$$\frac{d\varepsilon_i}{dt} = -\lambda\varepsilon_i + cv_i p_i, \quad i = 1, 2. \tag{12}$$

The differential equation for the total number of individuals $N = N_1 + N_2$ now becomes (Scheuring *et al.*, 2003):

$$\frac{dN}{dt} = -\kappa N + q(2 - D_0)vN^{-\beta}. \tag{13}$$

where

$$v \equiv p_1 v_1 + p_2 v_2 \tag{14}$$

is an average velocity, but note that it is not constant since p_i depend on the population numbers. (Since $\varepsilon_1/\varepsilon_2 = N_1/N_2 = z$, this relation is identical to (10). In Eq. 13 q is, again, a geometric constant. Using (12) and the fact that, due to the fractality, $\epsilon_i = \epsilon_0^{2(\beta+1)} N_i N^{\beta}$, $i = 1, 2$, (Scheuring *et al.*, 2003) the set of coupled population dynamical equations for the two competing replicators can be derived as

$$\frac{dN_i}{dt} = -\kappa N_i - q(D_0 - 1)vN^{-\beta-1}N_i + q v_i p_i \,(N_1/N_2)\, N^{-\beta}, \tag{15}$$

$i = 1, 2$, with $N = N_1 + N_2$, and p_i depending on N_1/N_2. By summing over i in (15) one recovers (13). Expression (15) represents a strongly coupled set of nonlinear equations with a novel type of power-law behaviour (with negative exponent $-\beta$). This set of equations can be considered as a population dynamics describing the coupling of two populations *mixing* on a fractal, and as we show below, opens up the possibility to have a nontrivial coexistence.

The steady-state coexistence analysis boils down to the study of the existence and stability of the fixed points for the population dynamics equations above. A straightforward linear stability analysis for this system around the coexistence (non-trivial) fixed points shows that the linear stability matrix has two eigenvalues (Scheuring *et al.*, 2003):

$$\Lambda_+ = -\lambda(1 - \alpha), \quad \Lambda_- = -\lambda. \tag{16}$$

One eigenvalue of the width dynamics is always the negative of the chaotic advection's positive Lyapunov exponent. As long as the parameter α is less than unity, the other eigenvalue is also negative. Thus it follows that for $0 < \alpha < 1$ *coexistence is stable*, and for $\alpha > 1$ it becomes *unstable*.

The presented mathematical forms and the conditions for coexistence remain valid if more than two species live in open chaotic flow, a numerical evidence for

which has been reported by Károlyi *et al.* (2000). It is natural to expect that the probabilities p_i, $i = 1, ..., m$ appear in the generalized form of $p_i = (\omega_i \varepsilon_i^\alpha)/(\sum_{i=1}^m \omega_i \varepsilon_i^\alpha)$, where ε_i are the partial width of the species and ω_i are phenomenological constants.

References

Allee, W. C.: 1938, *The Social Life of Animals*, Norton, New York.

Aref, H.: 1984, Stirring by Chaotic Advection, *J. Fluid Mech.* **143**, 1–21.

Aref, H., Jones, S. W., Mofina, S. and Zawadski, I.: 1989, Vortices, Kinematics and Chaos, *Physica D* **37**, 423–440.

Arístegui, J., Tett, P., Hernández-Guerra, A., Basterretxea, G., Montero, M. F., Wild, K., Sangrá, P., Hernández-León, S., Cantón, M., García-Braun, J.A., Pacheco, M. and Barton, E. D.: 1997, The Influence of Island-Generated Eddies on Chlorophyll Distribution: A Study of Mesoscale Variation Around Gran Canaria, *Deep-See Research I* **44**, 71–96.

Barton, E. D., Arístegui, J., Tett, P., Cantón, M., García-Braun, J. A., Hernández-León, S., Nykjaer, L., Almeida, C., Almunia, J., Ballesteros, S., Basterretxea, G., Escánez, J., García-Weill, L., Hernández-Guerra, A., López-Laatzen, F., Molina, R., Montero, M.F., Navarro-Pérez, E., Rodríguez, J.M., van Lenning, K., Vélez, H. and Wild, K.: 1998, The Transition Zone of the Canary Current Upwelling Region, *Progress in Oceanography* **41**, 455–504.

Boerlijst, M. C. and Hogeweg, P.: 1991, Spiral Wave Structures in Prebiotic Evolution: Hypercycles Stable Against Parasites, *Physica D* **48**, 17–28.

Cronhjort, M. B. and Blomberg, C.: 1995, 'Hypercycles versus Parasites in a Two-Dimensional Partial Differential Equation Model, *J. Theor. Biol.* **169**, 31–49.

Chaiken, J., Chevray, R., Tabor, M. and Tan, Q. M.: 1986, Experimental Study of Lagrangian Turbulence in a Stokes Flow, *Proc. Roy. Soc. London A* **408**, 165–174.

Czárán, T. and Szathmáry, E.: 2000, Coexistance of Replicators in Prebiotic Evolution, in Dieckmann, U., Law, R. and Metz, J. A. J. (eds), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, Cambridge U.P., Cambridge.

Doudna, J. A. and Cech, T. R.: 2002, The Chemical Repertoire of Natural Ribozymes, *Nature* **418**, 222–228.

Eigen, M.: 1971, Self-Organization of Matter and the Evolution of Macromolecules, *Naturwissenschaften* **58**, 465–523.

Eigen, M., Schuster, P., Gardiner, W. and Winkler-Oswatitsch R.: 1981, The Origin of Genetic Information. *Sci. Am.* **244(4)**, 78–94.

Eigen, M. and Schuster, P.: 1979, *The Hypercycle*, Springer, Berlin.

Ekland, E. H. and Bartel, D. P.: 1996, RNA-Catalyzed RNA Polymeryzation Using Nucleoside Triphisphates, *Nature* **382**, 373–376.

Gánti, T.: 1979, *A Theory of Biochemical Supersystems and its Application to Problems of Natural and Artificial Biogenesis*, Akadémiai Kiadó, Budapest and University Park Press, Baltimore.

Gilbert, W.: 1986, The RNA World, *Nature* **319**, 618.

Giona, M., Adrover, A., Muzzio, F. J., Cerbelli, S. and Alvarez, M. M.: 1999, The Geometry of Mixing in Time-periodic Chaotic Flows, I. Asymptotic Directionality in Physically Realizable Flows and Global Invariant Properties, *Physica D* **132**, 298–324.

Grey, D., Hutson, V. and Szathmáry, E.: 1995, A Re-Examination of the Stochastic Corrector Model, *Proc. Roy. Soc. Lond. B* **262**, 29–35.

Holm, N. G.: 1992, Marine Hydrothermal Systems and the Origin of Life, *Origins Life Evol. Biosphere* **22**, 5–14.

James, K. D. and Ellington, A. D.: 1999, The Fidelity of Template-Directed Oligonucleotide Ligation and the Inevitability of Polymerase Function, *Orig. Life. Evol. Biosphere* **29**, 375–390.

Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. and Bartel, D. P.: 2001, RNA-Catalyzed RNA Polimerization: Accurate and General RNA-Templated Primer Extension, *Science* **292**, 1319-1325.

Joyce, G. F. and Orgel, L. E.: 1999, Prospects for Understanding the Origin of the RNA World, in Gesteland, R. F., Cech, T. R. and Atkins, J. F. (eds), *The RNA World*, pp. 49–77, Cold Spring Harbor Lab. Press, Cold Spring Harbor.

Jung, C., Tél, T. and Ziemniak, E.: 1993, Application of Scattering Chaos to Particle Transport in a Hydrodynamical Flow, *Chaos* **3**, 555–568.

Károlyi, G., Péntek, Á., Scheuring, I., Tél, T. and Toroczkai, Z.: 2000, Chaotic Flow: the Physics of Species Coexistence, *Proc. Nat. Acad. Sci*. **97**, 13661–13665.

Károlyi, G., Péntek, Á., Toroczkai, Z., Tél, T. and Grebogi, C.: 1999, Chemical or Biological Activity in Open Chaotic Flows, *Physical Review E* **59**, 5468–5481.

Károlyi, G., Scheuring, I. and Czárán, T.: 2002, Metabolic Network Dynamics in Open Chaotic Flow, *Chaos* **12**, 460–469.

Károlyi, G. and Tél, T.: 1997, Chaotic Tracer Scattering and Fractal Basin Boundaries in a Blinking Vortex-Sink System, *Physics Reports* **290**, 125–147.

Kauffman, S. A.: 1986, Autocatalytic sets of proteins, *J. Theor. Biol.* **119**, 1–24.

Koch, A. L.: 1985, Primeval Cells: Possible Energy-Generating and Cell-Division Mechanisms, *J. Molec. Evol*. **21**, 270–277.

Landweber, L. F.: 1999, Experimental RNA Evolution, *Trends Ecol. Evol*. **14**, 353–358.

Lifson, S. and Lifson, H.: 1999, Models of Prebiotic Replication: Survival of the Fittest versus Extinction of the Unfittest. *J. Theor. Biol.* **199**, 425-433.

Maynard Smith J.: 1979, Hypercycles and the Origin of Life, *Nature* **20**, 445–446.

Maynard Smith, J.: 1983, Models of Evolution, *Proc. Roy. Soc. Lond. B* **219**, 315–325.

Maynard Smith, J. and Szathmáry, E.: 1995, *The Major Transitions in Evolution*, Freeman, Spektrum, Oxford.

Metcalfe, J. and Ottino, J.M.: 1994, Autocatalytic Processes in Mixing Flows, *Phys. Rev. Lett*. **72**, 2875–2878.

Ottino, J. M.: 1989, *The Kinematics of Mixing: Stretching, Chaos and Transport*, Cambridge U.P., Cambridge.

Ottino, J. M.: 1990, Mixing, Chaotic Advection, And Turbulence, *Ann. Rev. Fluid Mech*. **22**, 207–253.

Péntek, Á., Károlyi, G., Scheuring, I., Tél, T., Toroczkai, Z., Kadtke, J. and Grebogi, C.: 1999, Fractality, Chaos, and Reactions in Imperfectly Mixed Open Hydrodynamical Flows, *Physica A* **274**, 120–131.

Péntek, Á., Tél, T. and Toroczkai, Z.: 1995a, Chaotic Advection in the Velocity Field of Leapfrogging Vortex Pair, *J. Phys. A* **28**, 2191–2216.

Péntek, Á., Tél, T. and Toroczkai, Z.: 1995b, Fractal Tracer Patterns in Open Hydrodynamical Flows: the Case of Leapfrogging Vortex Pairs, *Fractals* **3**, 33–53.

Poole A., Jeffares, D. and Penny, D.: 1999, Early Evolution: Prokaryotes, the New Kids on the Block, *BioEssays* **21**, 880–889.

Rom-Kedar, V., Leonard, A. and Wiggins, S.: 1990, An Analytical Study of Transport, Mixing and Chaos in an Unsteady Vortical Flow, *J. Fluid. Mech*. **214**, 347–394.

Rom-Kedar, V. and Poje, A. C.: 1999, Universal Properties of Chaotic Transport in the Presence of Diffusion, *Physics of Fluids* **11**, 2044–2057.

Scheuring, I.: 2000, Avoiding Catch-22 of Early Evolution by Stepwise Increase in Copying Fidelity, *Selection* **1–3**, 135–145.

Scheuring, I. and Szathmáry, E.: 2001, Survival of Replicators with Parabolic Growth Tendency and Exponential Decay, *J. Theor. Biol.* **212**, 99–105.

Scheuring, I., Károlyi, G., Toroczkai, Z., Tél, T. and Péntek, Á.: 2003, Competing Populations in Flows with Chaotic Mixing, *Theor. Popul. Biol.*. **63**, 77–90.

Segré, D., Ben-Eli, D. and Lancet, D.: 2001, The Lipid World, *Orig. Life Evol. Biosphere* **31**, 119–145.

Shapiro, R.: 1984, The Improbability of Prebiotic Nucleid Acid Synthesis, *Orig. Life Evol. Biosphere* **14**, 565–570.

Shariff, K., Pulliam, T. H. and Ottino, J. M.: 1991, A Dynamical Systems Analysis of Kinematics in the Time-periodic Wake of a Circular Cylinder, *Lect. Appl. Math*. **28**, 613–646.

Sievert, D. and von Kiedrowski, G.:1994, Self-replication of Complementary Nucleotide Based Oligomers, *Nature* **369**, 221–224.

Sommerer, J. C., Ku, H.-C. and Gilreath, H. E.: 1996, Experimental Evidence for Chaotic Scattering in a Fluid Wake, *Phys. Rev. Lett*. **77**, 5055–5058.

Sommerer, J. C. and Ott, E.: 1993, Particles Floating on a Moving Fluid: A Dynamically Comprehensible Physical Fractal, *Science* **259**, 335–339.

Stadler, B. M. R., Stadler, P. F. and Wills P. R.: 2002, Evolution in Systems of Ligation Based Replicators, *Z. Phys. Chem.* **216**, 21-34.

Szabó, P., Scheuring, I., Czárán, T. and Szathmáry, E.: 2002, In Silico Simulations Reveal that Replications with Limited Dispersal Evolve Towards Higher Efficiency and Fidelity, *Nature* **420**, 360–363.

Szathmáry, E.: 1989, The Integration of the Earliest Genetic Information, *Trends Ecol. Evol.* **4**, 200–204.

Szathmáry, E. and Demeter, A.: 1987, Group Selection of Early Replicators and the Origin of Life, *J. Theor. Biol*. **128**, 463–486.

Szathmáry, E. and Gladkih, I.: 1989, Coexistence of Non-Enzymatically Replicating Templates, *J. Theor. Biol*. **138**, 55–58.

Tél, T.: 1990, Transient Chaos, in Hao, B.-L. (ed.), *Directions in Chaos* Vol. 3, pp. 149–211, World Scientific, Singapore.

Tél, T., Károlyi, G., Péntek, Á., Scheuring, I., Toroczkai, Z., Grebogi, C. and Kadtke, J.: 2000, Fractality, Chaos, Reactions, and Population Dynamics in Imperfectly Mixed Open Hydrodynamical Flows, *Chaos* **10**, 89–99.

Toffoli, T. and Margolus, N.: 1987, *Cellular Automata Machines: A New Environment for Modeling*, MIT Press, Cambridge, Mass.

Toroczkai, Z., Károlyi, G., Péntek, Á. and Tél, T.: 2001, Autocatalytic Reactions in Systems with Hyperbolic Mixing: Exact Results for the Active Baker Map, *J. Physics A* **34**, 5215–5235.

Toroczkai, Z., Károlyi, G., Péntek, Á., Tél, T. and Grebogi, C.: 1998, Advection of Active Particles in Open Chaotic Flows, *Phys. Rev. Lett.* **80**, 500–503.

Toroczkai, Z., Károlyi, G., Péntek, Á., Tél, T., Grebogi, C. and Yorke, J. A.: 1997, Wada Dye Boundaries in Open Hydrodynamical Flows, *Physica A* **239**, 235–243.

von Kiedrowsky, G. and Szathmáry E.:2000, Selection versus Coexistence of Parabolic Replicators Spreading on Surfaces, *Selection* **1-3**, 173–179.

Wächtershäuser, G.: 1988, Before Enzymes and Templates: Theory of Surface Metabolism, *Microbiol. Rev*. **52**, 452–484.

Wills, P. R., Kauffman, S. A. and Stadler, B. M. R.: 1998, Selection Dynamics in Autocatalytic Systems: Templates Replicating through Binary Ligation, *Bull. Math. Biol.* **60**, 1073–1098.

Wilson, D. S.: 1980, *The Natural Selection of Populations and Communities*, Benjamin Cummings, Menlo Park, California.

Zaug, A. J. and Cech, T. R.: 1986, The Intervening Sequence RNA of *Tetrahymena* is an Enzyme, *Science* **231**, 470–475.

Ziemniak, E., Jung, C. and Tél, T.: 1994, Tracer Dynamics in Open Hydrodynamical Flows as Chaotic Scattering, *Physica D* **76**, 123–146.

Zintzaras, E., Mauro, S. and Szathmáry, E.: 2002, 'Living' under the Challenge of Information Decay: The Stochastic Corrector Model versus Hypercycles, accepted by *J. Theor. Biol*.