

# Evolution in Systems of Ligation-Based Replicators

BÄRBEL M. R. STADLER<sup>a</sup>, PETER F. STADLER<sup>a,b</sup>,  
AND PETER R. WILLS<sup>c,1</sup>

<sup>a</sup>Institute for Theoretical Chemistry and Molecular Structural Biology,  
University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria  
E-Mail: {baer,studla}@tbi.univie.ac.at  
URL: <http://www.tbi.univie.ac.at/~{baer,studla}>

<sup>b</sup>The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

<sup>c</sup>Biomolecular Information Processing, GMD German National Research Centre for  
Information Technology, Schloss Birlinghoven, St Augustin, D-53754 Germany  
<http://www.phy.auckland.ac.nz/Staff/prw/>

## Abstract

The population dynamics of macromolecules that reproduce by means of template-directed ligation of two fragments are shown to be represented by a replicator equation with a special non-linear response function. This result is obtained through detailed consideration of the mechanism of ligation autocatalysis. In contrast to treatments which involve simplification to a parabolic growth law and the expectation of global coexistence of all species, we find that strong selection can take place in such systems, even when there is slow uncatalysed synthesis of replicators. Also, systems of this type are subject to invasion by new species that have a selective advantage. An expression is derived for the survival threshold in terms of species parameters and it is shown that this threshold depends on the total concentration of all species in the system. For a plausible distribution of species parameters, the number of surviving species coexisting above the threshold increases monotonically with increasing concentration. Illustrative numerical simulations are presented.

---

<sup>1</sup>Address for correspondence: Peter R. Wills, Department of Physics, The University of Auckland, Private Bag 92019, Auckland, New Zealand; Phone: +64-9-373 7599 ext. 8889, Fax: +64-9-373 7445

**Keywords:**

Autocatalytic Reaction Networks, Parabolic Growth, Extinction Threshold, Metadynamics

**1. Introduction**

Non-enzymatic template-directed self-replication of a biopolymer was first demonstrated in the case of a hexanucleotide synthesized from two trinucleotides more than a decade ago [26]. The recent discovery of a peptide that can undergo autocatalytic synthesis from two roughly equal fragments [14] has prompted a renewed interest in the theory of "parabolic" autocatalytic growth patterns. Ligation-based replication mechanisms appear to be of importance in a prebiotic setting [19]. A recent experimental study using nucleic acids [13] emphasizes the importance of ligation-based replication mechanisms for the origin of life.

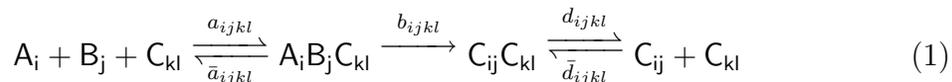
It is relatively easy to derive a rate equation displaying elementary parabolic growth behavior if one assumes that catalysis proceeds through the complementary binding of reactant(s) to free template and that autocatalysis is limited by the tendency of the template to bind to itself as an inactive "product inhibited" dimer [27]. However, in achieving an understanding of what is likely to happen in systems where there is a diverse mixture of reactants and catalytic templates, it is desirable to develop a detailed kinetic description of as many steps in the chemical process of template synthesis as is feasible and tractable from the mathematical point of view.

Szathmáry and Gladkih [24] simplified the dynamics of catalytic synthesis to a parabolic growth law  $\dot{x}_k \propto x_k^p$ ,  $0 < p < 1$  for the concentrations of the interacting template species. The resulting model suffers from a conceptual and a technical problem: (1) Under no circumstances does one observe extinction of a species for any parabolic growth law, and (2) the vector fields are not Lipschitz-continuous on the boundary of the concentration simplex, indicating that we cannot expect the model to give a physically reasonable description of the behavior in these regions.

In this contribution we analyze a more realistic version of the "parabolic growth law" that fixes both of these problems. Not surprisingly, we find that the resulting dynamics can be represented as a replicator equation with a monotonically decreasing response function. In the final section we briefly investigate a metadynamic model that considers the introduction of new species (either *mutants* related to members of the network or *immigrants* unrelated to the extant network members).

**2. Template Directed Ligation**

In a previous paper [28] we have derived the kinetic equations of a system of coupled template-instructed ligation reactions of the form



By assuming that the concentrations of the intermediates are stationary (the biochemical Michaelis-Menten *Ansatz*), we found

$$\frac{d}{dt}[\mathbf{C}_{kl}] = \sum_{ij} \tilde{m}_{kl ij} [\mathbf{C}_{ij}] \quad \text{with} \quad \tilde{m}_{kl ij} = [\mathbf{A}_k][\mathbf{B}_l] \frac{b_{kl ij} a_{kl ij}}{b_{kl ij} + \bar{a}_{kl ij}} \quad (2)$$

This equation for the concentration of the free templates is easily solved in the absence of fluxes. In most cases, however, one is more interested in the time evolution of the total concentration of the template molecules in the presence of a flux. It is convenient therefore to rewrite Eq.(2) in terms of the total template concentrations

$$c_{kl} = [\mathbf{C}_{kl}] \left\{ 1 + \sum_{ij} \frac{\tilde{m}_{ij kl}}{b_{ij kl}} + \sum_{ij} \frac{\tilde{m}_{ij kl}}{d_{ij kl}} \right\} + \sum_{ij \neq kl} \frac{\bar{d}_{ij kl}}{d_{ij kl}} [\mathbf{C}_{ij}] [\mathbf{C}_{kl}] + 2 \frac{\bar{d}_{kl kl}}{d_{kl kl}} [\mathbf{C}_{kl}]^2 + \sum_{ij} \frac{\tilde{m}_{kl ij}}{d_{ij kl}} [\mathbf{C}_{ij}]. \quad (3)$$

In this contribution we assume that the hetero-complexes  $\mathbf{C}_{kl} \mathbf{C}_{ij}$ ,  $ij \neq kl$ , can be neglected. In this case Eq.(3) simplifies considerably:

$$c_{kl} = [\mathbf{C}_{kl}] \left\{ 1 + \sum_{ij} \frac{\tilde{m}_{ij kl}}{b_{ij kl}} + 2 \frac{\tilde{m}_{kl kl}}{d_{kl kl}} \right\} + 2 \frac{\bar{d}_{kl kl}}{d_{kl kl}} [\mathbf{C}_{kl}]^2 \quad (4)$$

A closer inspection of Eq.(4) shows that we may drastically simplify the notation: (a) It is sufficient to use a single index to identify the individual species, hence we set  $kl \rightarrow k$  and  $ij \rightarrow i$ . We write  $\mathbf{A}_{i(k)}$  and  $\mathbf{B}_{j(k)}$  for the building blocks of  $\mathbf{C}_k = \mathbf{C}_{i(k)j(k)}$  (b) For each template species there are only two independent parameters determining the effective kinetics, namely

$$L_k = 1 + 2 \frac{\tilde{m}_{kk}}{d_{kk}} + \sum_i \frac{\tilde{m}_{ik}}{b_{ik}} \quad \text{and} \quad Q_k = 2 \frac{\bar{d}_{kk}}{d_{kk}} \quad (5)$$

Hence Eq.(3) is reduced to the simple quadratic equation,  $c_k = L_k [\mathbf{C}_k] + Q_k [\mathbf{C}_k]^2$  relating the total concentration to the free concentration of template-species  $C_k$ . As usual we introduce the total concentration  $c = \sum_j c_j$  of the template molecules and express concentration variables as relative abundances,  $x_k = c_k/c$  and  $y_k = [\mathbf{C}_k]/c$ . The relation of  $x_k$  and  $y_k$  becomes  $x_k = L_k y_k + c Q_k y_k^2$  in this notation. Solving for  $y_k$  we get

$$y_k = \frac{L_k}{2cQ_k} \left[ \sqrt{\frac{4cQ_k}{L_k^2} x_k + 1} - 1 \right] \quad (6)$$

In [28] we also translated the differential equations (2) into a system of ODEs in terms of total concentrations. Using relative abundances and neglecting all terms

pertaining to hetero-complexes, Eq.(44) of [28] reads

$$\dot{x}_k = \sum_i \tilde{m}_{ki} y_i - x_k \sum_{i,j} \tilde{m}_{ij} y_j. \quad (7)$$

Substituting Eq.(6) yields a system of differential equations in terms of relative concentrations only. The most interesting case of Eq.(7) is obtained by assuming *correct instruction*. In this case the matrix  $\tilde{M}$  is diagonal. We may then simplify the notation further by defining  $\beta_k = 4cQ_k L_k^{-2}$ ,  $\alpha_k = 2\tilde{m}_{kk}/L_k$  and introducing the non-linear function

$$\varphi(z) = \frac{1}{z}(\sqrt{z+1} - 1), \quad \varphi(0) = \frac{1}{2}. \quad (8)$$

The effective kinetic constants  $\alpha_k$  and  $\beta_k$  can be expressed in terms of the physical parameters as follows:

$$\begin{aligned} \alpha_k &= \frac{2[\mathbf{A}_{i(k)}][\mathbf{B}_{j(k)}]a_{kk}b_{kk}d_{kk}}{d_{kk}(b_{kk} + \bar{a}_{kk}) + [\mathbf{A}_{i(k)}][\mathbf{B}_{j(k)}]a_{kk}(2b_{kk} + d_{kk})} \\ \beta_k &= \frac{8d_{kk}\bar{d}_{kk}(b_{kk} + \bar{a}_{kk})^2}{(d_{kk}(b_{kk} + \bar{a}_{kk}) + [\mathbf{A}_{i(k)}][\mathbf{B}_{j(k)}]a_{kk}(2b_{kk} + d_{kk}))^2} c = \tilde{\beta}_k c \end{aligned} \quad (9)$$

We remark that the definition of  $\alpha_k$  above differs from the definition in [28, Eqs.(B3,B4)], while  $\beta_k$  is used in the same way. Our parameter  $\alpha_k$  equals  $\alpha_k\beta_k$  in the earlier work.

Note that the parameter  $\alpha_k$  is independent of the total concentration of templates while  $\beta_k$  is proportional to  $c$ . Eq.(7) now reads explicitly

$$\dot{x}_k = x_k \left( \alpha_k \varphi(\beta_k x_k) - \sum_j \alpha_j x_j \varphi(\beta_j x_j) \right) \quad (10)$$

This is a special form of a replicator equation with the non-linear response functions  $f_k(x) := \alpha_k \varphi(\beta_k x_k)$ . Eq.(10) is the subject of this investigation. If  $\tilde{M}$  is not diagonal but dominated by the diagonal entries we may regard Eq.(7) as being composed of a ‘‘correct instruction’’ model of the form (10) and mutation-type perturbation in the sense of [22], see also [28, sect.6].

Hence a growth law of the form  $\dot{x} \propto x\varphi(\beta x)$  replaces the troubling  $\dot{x} \propto \sqrt{x}$  in our more detailed analysis [28]. We formally recover Szathmary’s equation [24] by setting  $\varphi(z) \rightarrow z^{-1/2}$ . A brief inspection show that  $\varphi$  is continuously differentiable for all  $z > -1$ , i.e., it does not have the mathematical problems associated with the square root function. The first terms of the Taylor expansion around 0 are

$$\varphi(z) = \frac{1}{2} - \frac{1}{8}z + \frac{1}{16}z^2 - \mathcal{O}(z^3). \quad (11)$$

This expansion is also useful for evaluating  $\varphi$  numerically if  $|x|$  is very small. It is easily verified that the response function  $\varphi(z)$  is monotonously decreasing for  $z > -1$ .

### 3. Competition and Cooperation

Eq.(10) is a special case of a class of the replicator equations studied in [10]. Before we proceed to the general case, we consider the two-species dynamics in some detail. It is convenient to set  $x_1 = x$  and  $x_2 = 1 - x$ . It is sufficient then to consider only the equation for  $\dot{x} = \dot{x}_1$ . We have explicitly

$$\dot{x} = x(1-x)[\alpha_1\varphi(\beta_1x) - \alpha_2\varphi(\beta_2(1-x))] = F(x) \quad (12)$$

In the light of the general results reported in [10] it becomes clear that the following analysis is valid for any decreasing function  $\varphi(z)$ .

**Proposition 1.** *Suppose  $\varphi(z)$  is a continuously differentiable, monotonically decreasing function for all  $z \geq 0$ . Define*

$$\gamma_{12} = \frac{\alpha_1\varphi(\beta_1)}{\alpha_2\varphi(0)} \quad \text{and} \quad \gamma_{21} = \frac{\alpha_2\varphi(\beta_2)}{\alpha_1\varphi(0)}. \quad (13)$$

*Then the two-species model (12) has the following generic phase portraits:*

- (i) *If  $\gamma_{12} > 1$  then  $x_1 = 1$  is stable,  $x_2 = 1$  is unstable, and there is no interior rest point.*
- (ii) *Conversely, if  $\gamma_{21} > 1$  then  $x_2 = 1$  is stable,  $x_1 = 1$  is unstable, and there is no interior rest point.*
- (iii) *If  $\gamma_{12} < 1$  and  $\gamma_{21} < 1$  both corners are unstable and there is a unique stable interior rest point  $\hat{x}$ .*

*Proof.* We may use the discussion in [28] or verify by simple direct computation that the transversal eigenvalues at the two corner fixed points  $x_1 = x = 1$  and  $x_2 = 1$  (i.e.,  $x = 0$ ) are  $F'(1) = \alpha_2\varphi(0) - \alpha_1\varphi(\beta_1)$  and  $F'(0) = \alpha_1\varphi(0) - \alpha_2\varphi(\beta_2)$ , respectively.

(i) Assume that  $x_1 = 1$  is stable, i.e.,  $F'(1) < 0$ . Thus  $\alpha_1\varphi(\beta_1) > \alpha_2\varphi(0)$ . Monotonicity of  $\varphi$  implies now  $\alpha_1\varphi(0) > \alpha_1\varphi(\beta_1)$  and  $\alpha_2\varphi(0) > \alpha_2\varphi(\beta_2)$ ,  $\alpha_1\varphi(0) - \alpha_2\varphi(\beta_2) = F'(0) > 0$ . Thus  $x_2 = 1$  is unstable whenever  $x_1 = 1$  is stable. Returning to the definition of  $F(x)$  we find that stability of  $x_1 = 1$  implies

$$\alpha_1\varphi(0) \geq \alpha_1\varphi(\beta_1x) \geq \alpha_1\varphi(\beta_1) > \alpha_2\varphi(0) \geq \alpha_2\varphi(\beta_2(1-x)) \geq \alpha_2\varphi(\beta_2)$$

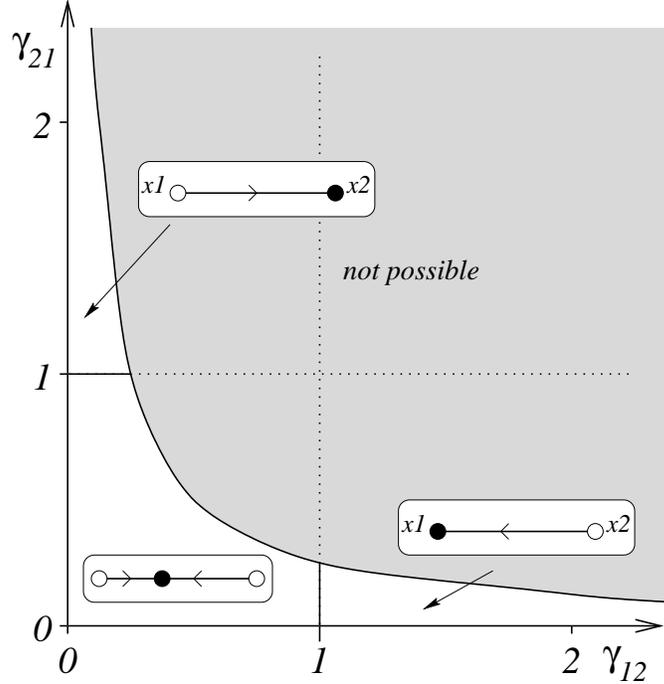
for all  $x \in [0, 1]$ , and hence  $\alpha_1\varphi(\beta_1x) - \alpha_2\varphi(\beta_2(1-x))$  does not have a zero in the physically relevant range  $[0, 1]$ .

Case (ii) is obtained from (i) by exchanging  $x_1$  and  $x_2$ .

(iii) Since the stability of  $x_1 = 1$  implies instability of  $x_2 = 1$ , and *vice versa*, in the remaining case both corners are unstable. Consider

$$g(x) = \alpha_1\varphi(\beta_1x) - \alpha_2\varphi(\beta_2(1-x)),$$

i.e., we write Eq.(12) as  $\dot{x} = x(1-x)g(x)$ . The zeros of  $g$  are determined by the intersection of the decreasing function  $\alpha_1\varphi(\beta_1x)$  and the increasing function  $\alpha_2\varphi(\beta_2(1-x))$ . Thus  $g$  has at most one zero  $\hat{x}$ . We have  $g(0) > 1$  and  $g(1) < 0$ . Continuity implies



**Figure 1.** Bifurcation diagram of the the general two-species model (12). The inaccessible part of the figure is drawn for  $\varphi(0) = 1/2$ , corresponding to Eq.(8).

that  $g$  has at least one zero in  $(0, 1)$ . The unique interior fixed point  $\hat{x}$  is necessarily stable because both corner equilibria are unstable.  $\square$

A schematic phase diagram for the two-species model is shown as Figure 1. Note that we have the general relation  $\gamma_{12}\gamma_{21} = \varphi(\beta_1)\varphi(\beta_2)$ , i.e., the parameters always satisfy  $0 < \gamma_{12}\gamma_{21} < \varphi(0)^2$ . Hence there is an excluded region in the phase diagram.

Let us now turn to the general case. Restating Theorem A from [10] in our notation immediately yields

**Proposition 2.** (i) *There is a unique fixed point  $\hat{x}$  which is the  $\omega$ -limit of all orbits in the interior of the simplex  $S_n$ .* (ii) *If  $\hat{x}$  lies in the interior of a face, then it is also the  $\omega$ -limit of all orbits in the interior of its face.* (iii) *If the species are labeled such that  $\alpha_1 \geq \alpha_2 \geq \dots \alpha_n$ , then there is an index  $m \geq 1$  such that  $\hat{x}$  is of the form  $\hat{x}_i > 0$  if  $i \leq m$  and  $\hat{x}_i = 0$  for  $i > m$ .* (iv) *If  $\min\{\beta_k\}$  is large enough, then  $m = n$ , i.e.,  $\hat{x}$  is an interior fixed point.*

**Remark.** It is shown in [11] that Eq.(10) is a Shashahani gradient system. Furthermore,  $V(x) = \prod_k x_k^{p_k}$ , where  $p_k$  are the coordinates of a globally stable fixed point is a Ljapunov function.

It is not hard to verify that the condition for survival of species  $k$ , item (iii) in Proposition 2, is explicitly

$$\alpha_k > 2\Phi(\hat{x}). \quad (14)$$

If we sort the replicating species according to decreasing values of the Darwinian fitness parameters,  $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$ , then there is an index  $m$  such that  $\hat{x}$  is of the form  $\hat{x}_k > 0$  if  $k \leq m$  and  $\hat{x}_k = 0$  for  $k > m$ . In other words,  $m$  species survive while the  $n - m$  least efficient replicators die out. This behavior is completely analogous to the reversible exponential competition case [18] where the rate constants  $a_k$  play the role of the Darwinian fitness parameters  $\alpha_k$ .

The concentration dependent values  $\beta_k$ , on the other hand, collectively influence the flux term and hence set the *extinction threshold*  $\alpha^* = 2\Phi(\hat{x})$ . The value of  $\alpha^*$  can be computed explicitly starting from the equation  $\alpha_k \varphi(\beta_k x_k) = \Phi(\hat{x})$  by means of the inverse function  $\varphi^{-1}(z) = (1 - 2z)/z^2$  of  $\varphi$ . We find

$$\hat{x}_k = \frac{1}{\Phi(\hat{x})^2} \frac{\alpha_k^2}{\beta_k} \left( 1 - 2 \frac{\Phi}{\alpha_k} \right) \quad (15)$$

Using the constraint  $\sum_k x_k = 1$ , we obtain

$$\Phi(\hat{x}) = \sum_k \frac{\alpha_k}{\beta_k} \left[ \sqrt{1 + \left( \sum_k \frac{\alpha_k^2}{\beta_k} \right) / \left( \sum_k \frac{\alpha_k}{\beta_k} \right)^2} - 1 \right] \quad (16)$$

where the sum runs over the surviving species only. Using the definition  $\tilde{\beta}_k = \beta_k/c$  from Eq.(9), we set

$$U_\ell = \left( \sum_{k=1}^{\ell} \frac{\alpha_k^2}{\tilde{\beta}_k} \right) / \left( \sum_{k=1}^{\ell} \frac{\alpha_k}{\tilde{\beta}_k} \right)^2 \quad \text{and} \quad V_\ell = \sum_{k=1}^{\ell} \frac{\alpha_k}{\tilde{\beta}_k} \quad (17)$$

For given  $\ell$  we can compute the threshold for the survival of the first  $\ell$  species as

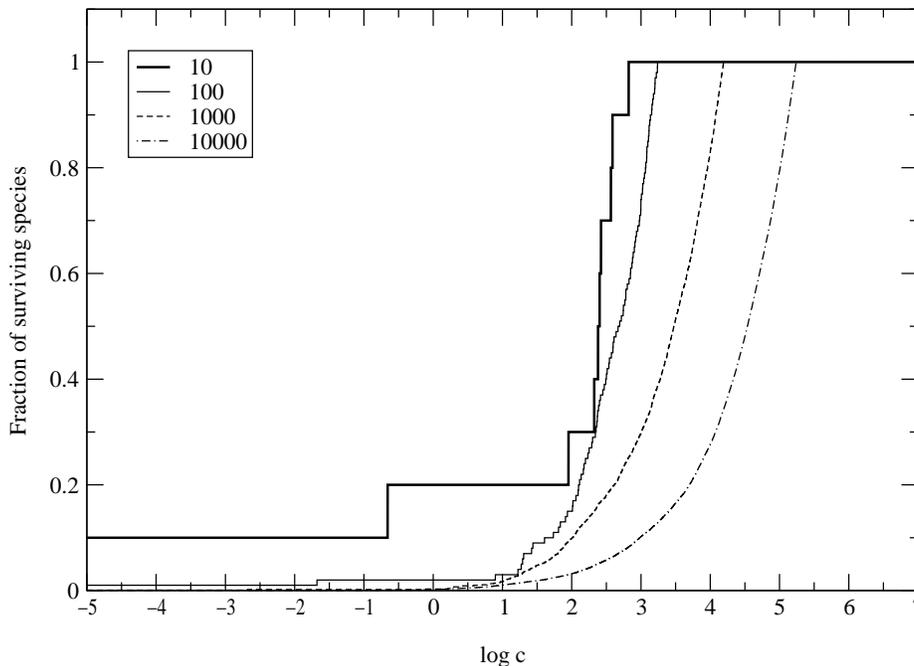
$$\Phi_\ell = c^{-1} V_\ell (\sqrt{1 + c U_\ell} - 1) = U_\ell V_\ell \varphi(c U_\ell) \quad (18)$$

For large  $c$  we find  $\Phi_\ell \sim c^{-1/2} \rightarrow 0$ . Thus  $\alpha_\ell > \alpha^*$  for all  $\ell$ , i.e., all species survive. As  $c \rightarrow 0$ , on the other hand  $\Phi_\ell \rightarrow U_\ell V_\ell / 2$  from below. The term

$$\bar{\alpha}_\ell = U_\ell V_\ell = \sum_{k=1}^{\ell} \alpha_k w_k / \sum_{k=1}^{\ell} w_k \quad \text{with} \quad w_k = \alpha_k / \tilde{\beta}_k \quad (19)$$

is a weighted average of the Darwinian fitness parameters of the survivors; in particular  $U_1 V_1 = \alpha_1$ . Hence  $\bar{\alpha}_\ell \leq \alpha_1$  and strictly decreasing with  $\ell$ . Thus only the fittest species survives for small enough  $c$ .

It is clear from Eq.(18) that  $\Phi_\ell$  decreases strictly monotonically with  $c$  because  $\bar{\alpha}_\ell > 0$  and  $U_\ell > 0$  are independent of  $c$ , and  $\varphi$  is a strictly monotonically decreasing



**Figure 2.** Fraction of surviving species as a function of the total concentration  $c$  for  $n = 10, 100, 1000, 10000$  species with  $\alpha_k \sim \exp(-k/n)$  and  $\tilde{\beta}_k$  chosen from Uniform $[0, 1]$ .

function. The fraction of surviving species thus increases monotonically with the total concentration, Fig. 2.

It is interesting to note that the Darwinian fitness parameters  $\alpha_k$  determine the order in which species reach extinction whereas the concentration dependent values  $\beta_k$  collectively influence the flux term and hence set the “extinction threshold”. In contrast to Szathmáry’s model equation [24] the extended replicator kinetics leads to both competitive selection and coexistence of replicators depending on total concentration and kinetic constants.

#### 4. Background Reactions

The Michaelis-Menten-type reaction mechanism (1) is likely to be accompanied by a slow background reaction of the form



Including (20) in the kinetic equations and simplifying the indices adds a term

$$\mathcal{B}_k(x) = \gamma_k - x_k \sum_l \gamma_l \quad (21)$$

where  $\gamma_k = g_{i(k)j(k)}[A_{i(k)}][B_{j(k)}]$ , see also [28]. This background term satisfies  $x_k = 0 \implies \mathcal{B}_k(x) > 0$  everywhere on the boundary of the state space, i.e., it satisfies the

conditions of a “mutation field” as defined in [22]. We can hence apply perturbation theory and in particular the *rest point migration theorem* to determine the qualitative effect of (20) as long as it is sufficiently slow compared to (1). Recall that a fixed point is *saturated* if it is stable perpendicular to the boundary of the state space [8, 9]. In biological terms that means that the equilibrium is uninvadable. In particular, every stable fixed point is saturated. The rest point migration theorem describes the behavior of the fixed points of a replicator equation with a small mutational perturbation: Saturated fixed points move into the interior of the state space, while non-saturated fixed points move into the non-physical exterior of the state space. The dynamics of the ligation based replication dynamics therefore remains qualitatively unchanged. The effect of the background reaction is simply to “pollute” the equilibrium with very small concentrations of those species that would have died out in its absence.

## 5. Evolution of Ligation-Replication Systems

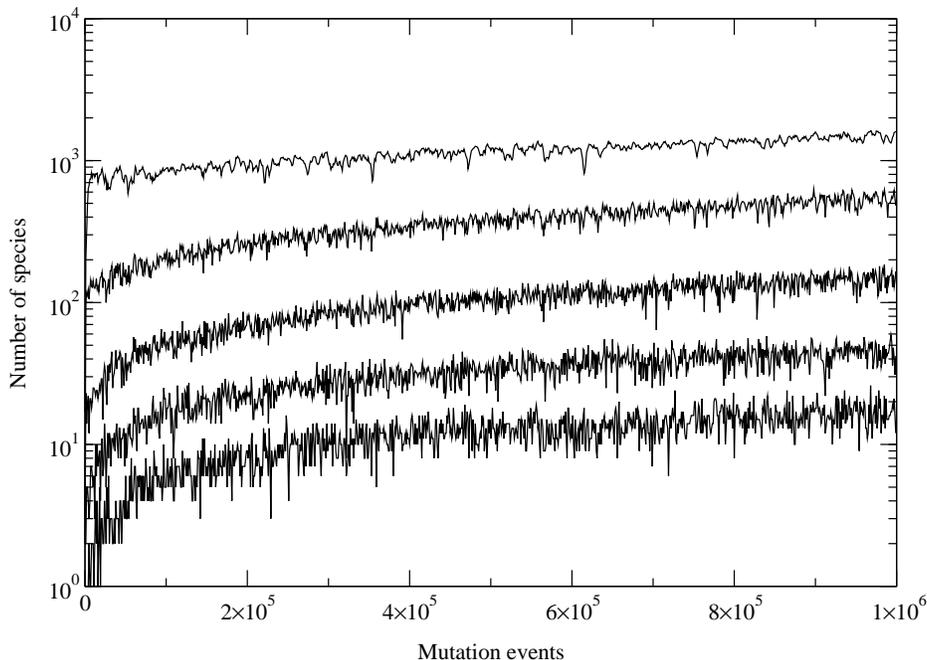
A species  $k$  is represented merely by the pair  $(\alpha_k, \beta_k)$  in our model. Detailed computer simulations of quasi-species-like evolution of RNA sequences have shown that the behavior of an evolving population depends crucially on the structure of the fitness landscape [6, 17, 12]. The reason is that periods of “stasis” are determined by the waiting times for the occurrence of advantageous mutants. A thorough discussion of this topic in terms of accessibility relations of phenotypes can be found in [4, 5, 21].

We take a much less ambitious point of view here. In order to identify trends in the evolution of the ligation-replication model described here we consider mutations of the form

$$(\alpha_k, \beta_k) \rightarrow (\alpha_k + \epsilon_1, \beta_k + \epsilon_2)$$

where  $\epsilon_1$  and  $\epsilon_2$  are independent random numbers drawn from a given distribution. Note that at this level of description we do not have an absolute time scale. Rather we measure “time” by the number of mutation events. If we allow for large mutations, i.e.,  $\mathbb{E}[|\epsilon_k|]$  and  $\alpha_k$  are of the same order of magnitude, then we occasionally observe a single surviving dominant mutant. Hence we only discuss the case of small mutational effects further.

Similar metadynamic models have been considered in the past for Lotka-Volterra models in the context of super-infection in host-parasite associations [16] or competition and biodiversity in spatially structured habitats [25]. The evolution of diversity in general second order replicator networks is considered in [7]. One finds a slow increase in the number of coexisting species with time provided the mutants are similar to the parents, while unrelated immigrants occasionally destroy the networks completely. Although coexistence is established in a completely different way in these networks, namely by means of positive feedback instead of product buffering, we find similar metadynamics.



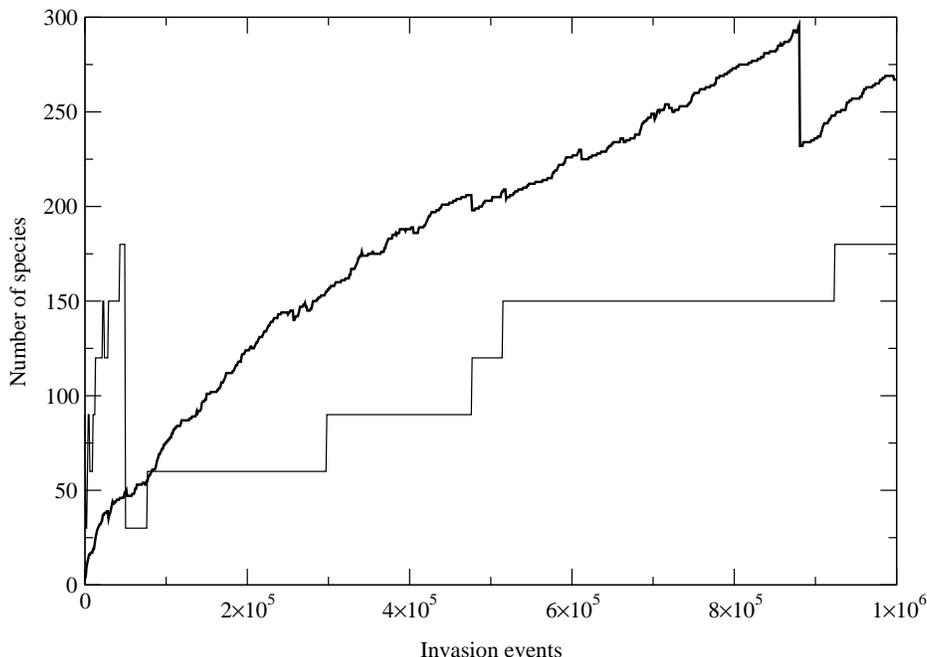
**Figure 3.** Number of species in a ligation-replication system with mutation. From top to bottom:  $c = 10, 1, 0.1, 0.01, 0.001$ . The simulation is started with  $\alpha_k \sim \exp(-\lambda k)$  and  $\beta_k$  taken from a uniform distribution.

In the case of mutation the number of surviving species slowly increases with time. The diversity of the ecosystem depends strongly on the total concentration  $c$ , see Fig. 3.

Simulations show no selection for particular values of the product-inhibition parameters  $\tilde{\beta}_k$ . This may seem surprising since the extinction threshold is dependent on the  $\beta$ -parameters. However the influence of species  $k$ 's value of  $\beta_k$  on the threshold value is very small, hence survival depends essentially on the value of the Darwinian fitness parameter  $\alpha_k$  relative to the externally determined threshold value.

It is interesting to note that the spread of the fitness distribution, i.e.,  $\alpha_1 - 2\Phi(\hat{x})$ , is comparable to the expected value  $\mathbb{E}[|\epsilon_k|]$  of the mutation effect.

Instead of mutation we might also consider invasion, i.e., at each time step a randomly generated invader  $(\alpha, \beta)$  is created. When the value of  $\alpha$  is taken from Uniform[0, 1] then the threshold  $2\Phi(\hat{x})$  approaches 1. The number of species continues to grow as additional invaders accumulate in the slowly narrowing gap between  $2\Phi(\hat{x})$  and 1. In contrast, when the invaders have a Gaussian distribution of  $\alpha_k$ , there is no systematic increase in the number of species. Occasionally species with a large selective advantage occur leading to the extinction of all their competitors, Fig. 4.



**Figure 4.** The number of surviving species under invasion. Thick line: uniform distribution of  $\alpha_k$ , thin line: Gaussian distribution with mean and standard deviation 0.5, truncated at 0. The values are multiplied by 30.

## 6. Discussion

By using the biochemical Michaelis-Menten *Ansatz* we have investigated the detailed mechanism of template-directed macromolecular ligation and derived a replicator equation describing this form of autocatalysis. These considerations remove certain difficulties inherent in more approximate treatments that result in a parabolic growth law and the prediction of global coexistence of all templates in a homogeneous system [24, 27].

In contrast, the more detailed approach shows that the possible behavior of these systems ranges from selection of the single fittest species, reminiscent of the *master sequence* in the quasi-species theory [1, 2], through coexistence of a proportion of species with fitnesses above some threshold, to the global coexistence of all species, as is predicted from the approximate parabolic growth law. This existence of an extinction threshold appears to be a generic feature of replication systems with product inhibition as it generalizes to higher order autocatalysis [20]. We note that the concentration dependence of the selection behavior of the system shows that it is not generally possible to reach conclusions about the generic behavior of a system just by studying the mode of selection at low concentration.

Having given a limited demonstration of this range of behaviors previously [28], we have now provided a rigorous, general, theoretical analysis of the phenomena and

we can conclude that systems of competing species replicating via template-directed ligation can meet two of the criteria requisite for true evolution: (i) strong selection leading to the extinction of some or even almost all species, and (ii) susceptibility to invasion by new, advantageous species.

This conclusion bears directly on the possibility of observing true selection in systems of replicators which reproduce through ligation autocatalysis [14, 26, 15]. The current theoretical treatment is directly relevant in the case that template complementarity is strong enough that cross reactions due to hetero-duplex formation ( $ij \neq kl$  in Eqs.(1-3)) are of no consequence. However, more important in the experimental context is the conclusion that strong selection can be expected to take place in the presence of slow background reactions (see Section 4). Although slow background reactions will produce small concentrations of all possible species  $\{C_{ij}\}$  that can arise from the available reactants  $\{A_i\}$  and  $\{B_j\}$ , this will have only a minor perturbative effect on the survival threshold, except in the case of an invading template whose kinetic parameters place it above the threshold, in which case there will be a metadynamic shift in the selection equilibrium. Thus, our analysis predicts that if the chemical kinetic parameters governing replication in experimental systems can be brought within appropriate ranges, perhaps through the choice of special conditions of pH, ionic strength, temperature, *etc.*, then strong selection of some species should be observed, even though the rate of growth (Eq.10) is, strictly speaking, sub-exponential.

The origin of life requires a mechanism of chemical replication in which strong selection enables some species to outgrow others, the “losers” which die out. On the other hand, the coexistence of more than one “master sequence” is required for functional specialisation and cooperation to emerge. Replicators which reproduce through duplex formation utilizing a Michaelis-Menten type of mechanism fulfil these conditions much more easily than replicators needing the strong and specific catalytic interactions characteristic of hypercyclic cooperation [3, 23]. Therefore they are good candidates for the first molecules which may have been selectively amplified in the prebiotic environment.

### Acknowledgements

We dedicate this paper to Peter Schuster in celebration of his 60th birthday. BMRS is supported by the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, Project No. P-13887-MOB. PRW is grateful to the Alexander von Humboldt Foundation for the provision of a research fellowship.

### References

- [1] M. Eigen. Selforganization of matter and the evolution of macromolecules. *Naturwiss.*, 58:465–523, 1971.
- [2] M. Eigen, J. S. McCaskill, and P. Schuster. The molecular quasi-species. *Adv. Chem. Phys.*, 75:149–263, 1989.

- [3] M. Eigen and P. Schuster. *The Hypercycle*. Springer-Verlag, New York, Berlin, 1979.
- [4] W. Fontana and P. Schuster. Continuity in Evolution: On the Nature of Transitions. *Science*, 280:1451–1455, 1998.
- [5] W. Fontana and P. Schuster. Shaping Space: The Possible and the Attainable in RNA Genotype-Phenotype Mapping. *J. Theor. Biol.*, 194:491–515, 1998.
- [6] W. Fontana, P. F. Stadler, E. G. Bornberg-Bauer, T. Griesmacher, I. L. Hofacker, M. Tacker, P. Tarazona, E. D. Weinberger, and P. Schuster. RNA folding landscapes and combinatorial landscapes. *Phys. Rev. E*, 47:2083–2099, 1993.
- [7] R. Happel and P. F. Stadler. The evolution of diversity in replicator networks. *J. Theor. Biol.*, 195:329–338, 1998.
- [8] J. Hofbauer. Saturated equilibria, permanence, and stability for ecological systems. In T. G. Hallam, L. J. Gross, and S. A. Levin, editors, *Mathematical ecology*, Singapore, 1988. World Scientific. Proc. Autumn Course Res. Semin., Miramare- Trieste/Italy 1986.
- [9] J. Hofbauer. An index theorem for dissipative semiflows. *Rocky Mt. J. Math.*, 20:1017–1031, 1990.
- [10] J. Hofbauer, P. Schuster, and K. Sigmund. Competition and cooperation in catalytic selfreplication. *J. Math. Biol.*, 11:155–168, 1981.
- [11] J. Hofbauer and K. Sigmund. *Dynamical Systems and the Theory of Evolution*. Cambridge University Press, Cambridge U.K., 1988.
- [12] M. A. Huynen, P. F. Stadler, and W. Fontana. Smoothness within ruggedness: the role of neutrality in adaptation. *Proc. Natl. Acad. Sci. (USA)*, 93:397–401, 1996.
- [13] K. D. James and A. D. Ellington. The fidelity of template-directed oligonucleotide ligation and the inevitability of polymerase function. *Orig. Life Evol. Biosph.*, 29:375–390, 1999.
- [14] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, and M. R. Ghadiri. A self-replicating peptide. *Nature*, 382:525–528, 1996.
- [15] S. Lifson and H. Lifson. A model for prebiotic replication: Survival of the fittest versus extinction of the unfittest. *J. Theor. Biol.*, 199:425–433, 1999.
- [16] R. M. May and M. A. Nowak. Superinfection, metapopulation dynamics, and the evolution of diversity. *J. Theor. Biol.*, 170:95–114, 1994.
- [17] P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker. From sequences to shapes and back: A case study in RNA secondary structures. *Proc. Roy. Soc. Lond. B*, 255:279–284, 1994.
- [18] P. Schuster and K. Sigmund. Dynamics of evolutionary optimization. *Ber. Bunsenges. Phys. Chem.*, 89:668–682, 1985.
- [19] P. Schuster and P. F. Stadler. Nature and evolution of early replicons. In E. Domingo, R. Webster, and J. Holland, editors, *Origin and Evolution of Viruses*, pages 1–24. Academic Press, London, UK, 1999.
- [20] B. M. R. Stadler, P. F. Stadler, and P. Schuster. Dynamics of autocatalytic replicator networks based on higher order ligation reactions. *Bull. Math. Biol.*, 62:1061–1086, 2000.
- [21] B. M. R. Stadler, P. F. Stadler, G. Wagner, and W. Fontana. The topology of the possible: Formal spaces underlying patterns of evolutionary change. *J. Theor. Biol.*, 2000. submitted, SFI preprint 00-12-070.
- [22] P. F. Stadler and P. Schuster. Mutation in autocatalytic networks - an analysis based on perturbation theory. *J. Math. Biol.*, 30:597–631, 1992.
- [23] P. F. Stadler and P. Schuster. Permanence of sparse autocatalytic networks. *Math. Biosc.*, 131:111–134, 1996.
- [24] E. Szathmary and I. Gladkih. Sub-exponential growth and coexistence of non-enzymatically replicating templates. *J. Theor. Biol.*, 138:55–58, 1989.

- [25] D. Tilman. Competition and biodiversity in spatially structured habitats. *Ecology*, 75:2–16, 1994.
- [26] G. von Kiedrowski. A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed. Engl.*, 25:932–935, 1986.
- [27] G. von Kiedrowski. Minimal replicator theory I: Parabolic versus exponential growth. In *Bioorganic Chemistry Frontiers, Volume 3*, pages 115–146, Berlin, Heidelberg, 1993. Springer-Verlag.
- [28] P. R. Wills, S. A. Kauffman, B. M. Stadler, and P. F. Stadler. Selection dynamics in autocatalytic systems: Templates replicating through binary ligation. *Bull. Math. Biol.*, 60:1073–1098, 1998.