Dynamics of Autocatalytic Replicator Networks Based on Higher Order Ligation Reactions

BÄRBEL M. R. STADLER
Institut für Theoretische Chemie und Molekulare Strukturbiologie,
Universität Wien, Vienna, Austria

PETER F. STADLER
Institut für Theoretische Chemie und Molekulare Strukturbiologie,
Universität Wien, Vienna, Austria†
The Santa Fe Institute, Santa Fe, New Mexico

PETER SCHUSTER
Institut für Theoretische Chemie und Molekulare Strukturbiologie,
Universität Wien, Vienna, Austria
The Santa Fe Institute, Santa Fe, New Mexico

E-mail: {baer,studla,pks}@tbi.univie.ac.at, or
{bmrs,stadler,pks}@santafe.edu

A class of autocatalytic reaction networks based on template dependent ligation and higher order catalysis is analyzed. Apart from an irreversible ligation reaction we consider only reversible aggregation steps that provide a realistic description of molecular recognition. The over-all dynamics can be understood by means of replicator equations with highly non-linear interaction functions. The dynamics depends crucially on the total concentration c_0 of replicating material.

For small c_0, in the hyperbolic growth regime, we recover the familiar dynamics of second order replicator equations with its wealth of complex dynamics ranging from multi-stability to periodic and strange attractors as well as to heteroclinic orbits. For large c_0, in the parabolic growth regime, product inhibition becomes dominating and we observe a single globally stable equilibrium tantamount to permanent coexistence. In an intermediate parameter range we sometimes observe a behavior that is reminiscent of "survival of the fittest". Independently replicating species (Schlögel's model) and the hypercycle are discussed in detail.

Key words: Autocatalytic Networks, Replicator Equation, Ligation, Hypercycle

†Address for correspondence:
Peter F. Stadler, Institut für Theoretische Chemie und Molekulare Strukturbiologie, Universität Wien Währingerstraße 17, A-1090 Wien, Austria Phone: **43 1 4277 52738, Fax: **43 1 4277 52793.

1. Introduction

1.1. *The Origin of Replication.* Most of the building blocks of present day biomolecules are available from different prebiotic sources, from extraterrestrial origins as well as from processes that took place in the primordial atmosphere or near hot vents in deep oceans [26]. Condensation reactions and polymerization reactions formed non-instructed polymers, such as random oligopeptides of the protenoid type [9].

Julius Rebek and his coworkers [39, 44] demonstrated that autocatalytic chemical reaction systems can be assembled from simple organic molecules that undergo self-replication. Autocatalysis in such systems of small molecules is certainly interesting because it reveals some mechanistic details of molecular recognition. However, these systems are hardly at the basis of biological replication since they cannot store a sizeable amount of information. Consequently, there is no room for molecular variants, "mutations", and hence these systems cannot sustain Darwinian evolution.

Ligation of small pieces to larger units, on the other hand, is a source of combinatorial complexity providing sufficient capacity for information storage and, hence, offers the possibility of evolution. Heteropolymer formation thus seems inevitable and we shall therefore focus on replicons which have this property: nucleic acids and proteins.

Although there are a number of difficulties concerning the origin of RNA molecules directly from an unorganized prebiotic chemistry, see e.g. [21, 29, 30, 35], the notion of an *RNA world* [10, 11, 21] is widely considered a quite plausible one: RNA is an obligatory template and hence capable of storing retrievable information, and it has catalytic properties. While the catalytic properties of RNA are less universal than those of proteins, they are apparently sufficient for processing RNA. A recent experimental study using nucleid acids [20] emphasizes the importance of ligation-based replication mechanisms for the origin of life.

The natural ribozymes discovered early on, such as the RNA moiety of RNase P [12], the class I introns [3] or the "hammerhead" [45], were all RNA cleaving molecules. Selection from (partially) random RNA libraries and subsequent evolutionary optimization, however, has revealed that their activity is by no means restricted to RNA cleavage: Some ribozymes show ligase activity [1, 8] and many efforts were undertaken to prepare a ribozyme with full RNA replicase activity. The attempt that comes closest to the goal yielded a ribozyme that catalyzes RNA polymerization in short stretches [7].

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 [46]. The experimental investigation of competition among non-enzymatically replicating nucleotide templates has been limited by the small size and restricted variability of the species among which selection specificity has been measured [38, 48].

More recently it was shown that a peptide comprised of 32 amino acids can undergo exactly the same form of autocatalytic synthesis from two roughly equal fragments [22]. In contrast to the obligatory template function of RNA, the "self-replicating peptides" discovered by Reza Ghadiri's group require a very special three dimensional structure [23, 24, 25, 51]. Hence the prerequisites of open-ended evolutions — conservation and propagation of mutants — might be very hard, if not impossible, to meet in such a system.

1.2. ***Networks of Replicating Molecules.*** It is relatively easy to derive a kinetic rate equation displaying the elementary behavior of replicons under the assumption that catalysis proceeds through 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 [47].

In order to achieve an understanding of what is likely to happen in systems containing a diverse mixture of reactants and catalytic templates, however, it is necessary to develop a comprehensive kinetic description. A model at the level of elementary chemical reactions is not only beyond present capacities but even if possible will not provide much insight because of the excessive number of kinetic constants that would have to be measured, computed, or at least seriously estimated. We suggest that the appropriate level of description is therefore the over-all kinetics of the replication process. As an example consider the following system of competing templates that replicate via template instructed but otherwise unassisted ligation:

$$\mathsf{A_i} + \mathsf{B_j} + \mathsf{C_{kl}} \underset{\bar{a}_{ijkl}}{\overset{a_{ijkl}}{\rightleftharpoons}} \mathsf{A_iB_jC_{kl}} \xrightarrow{b_{ijkl}} \mathsf{C_{ij}C_{kl}} \underset{\bar{d}_{ijkl}}{\overset{d_{ijkl}}{\rightleftharpoons}} \mathsf{C_{ij}} + \mathsf{C_{kl}}\,. \tag{1}$$

The reaction mechanism (1), which is analyzed in detail in [49], comprises the simplest possible model that explicitly takes into account that the complexes mediating molecular recognition have a significant — and sometimes determining — influence on the reaction kinetics. There are two limiting cases with very different dynamical behavior:

(i) In the limit of small total concentration and/or unstable dimers one obtains the quasispecies model. It is characterized by strong selection leading to "survival of the fittest". In the error prone case, it poses a severe limit on the amount of information that can be transferred into the next generation, requiring very accurate replication from the very beginning [5]. Selection in this case is too strong to explain the evolution of diverse ecologies.

(ii) For large concentrations and/or very stable template-template complexes $\mathsf{C_{ij}C_{kl}}$, one obtains a parabolic growth-law [43]. It leads to global coexistence of all reaction products. In the essentially infinite space opened by the combinatorics of hetero-polymers, such a system will be littered by all sorts of reaction products, and eventually be indistinguishable from a random pool of sequences. Selection in this case is too weak to explain the emergence of order.

In the intermediate regime we find an *extinction threshold* depending on the total concentration of replicating material such that all templates with replication rates below the threshold eventually vanish [34].

1.3. ***Higher Order Catalysis.*** Enzyme-free template-induced synthesis of longer RNA molecules from monomers, however, has not been successfully achieved so far (see e.g. [28]). A major problem, among others, is the dissociation of double stranded molecules at the temperature of efficient replication: If monomers bind with sufficiently high binding constants to the template in order to guarantee the desired accuracy of replication, then the new molecules are too sticky to dissociate after the synthesis has been completed. Reaction complexes involving more components than the template and the newly formed ligation product might facilitate the release of the template by binding the reaction product to another component.

A second problem arises from Eigen's *error threshold* [4, 5] which restricts the amount of information that can be sustained under error-prone replication. This limitation can be circumvent by assuming catalytic feedback whereby each template helps in the replication of the next one, thereby forming a *hypercycle*.

In this contribution we shall describe a plausible molecular mechanism for ligation based replication involving both a template and a second "catalyst". Experimentally such a system is at least reasonable since multiple helices are well known structures in both nucleic acids and coiled-coil peptides.

The generic features of the resulting reaction networks are largely unexplored despite some efforts towards understanding the dynamics of chemically realistic models of "prebiotic" replication in [15]. On the other hand, the limiting case of strong selection, in which the dynamics of complex formation and product inhibition plays no role, is very well understood. It is described by second order *replicator equations* [32]

$$\dot{x}_k = x_k \left(\sum_{j=1}^{n} a_{kj} x_j - \sum_{i,j} a_{ij} x_i x_j \right) . \tag{2}$$

which are the subject of a great number of research papers as well as of the book [18].

This contribution is organized as follows: In section 2 we describe the kinetic of a single replicating template, extending von Kiedrowski's *Minimal Replicator Theory* [47] to ternary template complexes, adding a regime of hyperbolic growth. The kinetic equations for competing species are derived in section 3. The limiting case of independently replicating templates, i.e., the ligation version of Schlögel's model, is analyzed in section 4. Then we consider more general aspects of "correct instruction" and investigate the ligation-hypercycle. The implications of our model are discussed in the final section 6.

2. A Single Template Species

We consider a template species C which can be obtained by the concatenation (ligation) of two building blocks A and B which are provided from the outside of the reaction vessel. The over-all replication reaction is therefore $\mathsf{A}+\mathsf{B}\rightarrow\mathsf{C}$. For simplicity we shall assume throughout this paper that the concentrations $[\mathsf{A}]$ and $[\mathsf{B}]$ of the building blocks are buffered, i.e., constant.

The crucial step of the replication mechanism is the formation of an intermediate species ABCC which we may visualize as a triple helix consisting of the template C, the building material A and B properly aligned along C so that the ligation of the parts A and B can readily proceed, and an additional copy of C that might be thought either as a specific catalyst that is involved in the ligation reaction, or its role might be seen in facilitating the release of the product C. In either case, we consider the following reaction scheme

$$
\begin{aligned}
\mathsf{A}+\mathsf{B}+\mathsf{C} &\underset{\bar{b}}{\overset{b}{\rightleftharpoons}} \mathsf{ABC}\\
\mathsf{A}+\mathsf{B}+\mathsf{CC} &\underset{\bar{f}}{\overset{f}{\rightleftharpoons}} \mathsf{ABCC}\\
\mathsf{C}+\mathsf{C} &\underset{\bar{d}}{\overset{d}{\rightleftharpoons}} \mathsf{CC}\\
\mathsf{ABC}+\mathsf{C} &\underset{\bar{h}}{\overset{h}{\rightleftharpoons}} \mathsf{ABCC}\\
\mathsf{CC}+\mathsf{C} &\underset{\bar{g}}{\overset{g}{\rightleftharpoons}} \mathsf{CCC}\\
\mathsf{ABCC} &\xrightarrow{a} \mathsf{CCC}
\end{aligned}
\tag{3}
$$

in a *continuously stirred tank reactor* (CSTR) with flux rate r. Note that apart from the first two steps, $\mathsf{A}+\mathsf{B}+\mathsf{C}\rightleftharpoons\mathsf{ABC}$ and $\mathsf{A}+\mathsf{B}+\mathsf{CC}\rightleftharpoons\mathsf{ABCC}$, all reversible reactions are bimolecular. While it would be possible to resolve the trimolecular steps into a sequence of bimolecular reactions, the trimolecular step is a good approximation provided the intermediates AC, BC, ACC, and BCC are short-lived, which we think is a very reasonable assumption.

The rules of mass action kinetics immediately lead to the following set of kinetic differential equations

$$\begin{aligned}
\frac{\mathrm{d}[\mathsf{ABC}]}{\mathrm{d}t} &= b[\mathsf{A}][\mathsf{B}][\mathsf{C}] - \bar{b}[\mathsf{ABC}] + \bar{h}[\mathsf{ABCC}] - h[\mathsf{ABC}][\mathsf{C}] - r[\mathsf{ABC}] \\
\frac{\mathrm{d}[\mathsf{ABCC}]}{\mathrm{d}t} &= f[\mathsf{A}][\mathsf{B}][\mathsf{CC}] - \bar{f}[\mathsf{ABCC}] + h[\mathsf{ABC}][\mathsf{C}] - \bar{h}[\mathsf{ABCC}] - a[\mathsf{ABCC}] - r[\mathsf{ABCC}] \\
\frac{\mathrm{d}[\mathsf{CC}]}{\mathrm{d}t} &= \bar{f}[\mathsf{ABCC}] - f[\mathsf{A}][\mathsf{B}][\mathsf{CC}] + d[\mathsf{C}][\mathsf{C}] - \bar{d}[\mathsf{CC}] + \bar{g}[\mathsf{CCC}] - g[\mathsf{CC}][\mathsf{C}] - r[\mathsf{CC}] \\
\frac{\mathrm{d}[\mathsf{CCC}]}{\mathrm{d}t} &= g[\mathsf{CC}][\mathsf{C}] - \bar{g}[\mathsf{CCC}] + a[\mathsf{ABCC}] - r[\mathsf{CCC}] \\
\frac{\mathrm{d}[\mathsf{C}]}{\mathrm{d}t} &= \bar{h}[\mathsf{ABCC}] - h[\mathsf{ABC}][\mathsf{C}] + 2\bar{d}[\mathsf{CC}] - 2d[\mathsf{C}][\mathsf{C}] + \bar{g}[\mathsf{CCC}] - g[\mathsf{CC}][\mathsf{C}] + \\
&\quad \bar{b}[\mathsf{ABC}] - b[\mathsf{A}][\mathsf{B}][\mathsf{C}] - r[\mathsf{C}]
\end{aligned} \tag{4}$$

We are interested in the behavior of the total concentration

$$c = [\mathsf{C}] + 2[\mathsf{CC}] + 3[\mathsf{CCC}] + [\mathsf{ABC}] + 2[\mathsf{ABCC}] \tag{5}$$

Assuming that the intermediates are stationary,

$$\frac{\mathrm{d}[\mathsf{ABC}]}{\mathrm{d}t} = \frac{\mathrm{d}[\mathsf{ABCC}]}{\mathrm{d}t} = \frac{\mathrm{d}[\mathsf{CC}]}{\mathrm{d}t} = \frac{\mathrm{d}[\mathsf{CCC}]}{\mathrm{d}t} = 0 \tag{6}$$

a lengthy computation, assisted by the use of `Mathematica` shows that

$$\frac{\mathrm{d}c}{\mathrm{d}t} = \frac{\mathrm{d}[\mathsf{C}]}{\mathrm{d}t} = a[\mathsf{ABCC}] - rc \tag{7}$$

Explicitly, the concentrations of free template $[\mathsf{C}]$ and total template c are related via

$$c = \frac{p_4[\mathsf{C}]^4 + p_3[\mathsf{C}]^3 + p_2[\mathsf{C}]^2 + p_1[\mathsf{C}]}{q_2[\mathsf{C}]^2 + q_1[\mathsf{C}] + q_0} \tag{8}$$

where $p_1, p_2, p_3, p_4, q_0, q_1, q_2 > 0$ and $q_2 = \mathcal{O}(r)$. The coefficients are very lengthy sums of products of the elementary rate constants defined in the reaction scheme (3), hence we refrain from displaying them. Explicitly, we have

$$\dot{c} = \frac{\mathrm{d}c}{\mathrm{d}t} = [\mathsf{A}][\mathsf{B}][\mathsf{C}]^2 \frac{\alpha + \gamma[\mathsf{C}]}{\beta + \delta[\mathsf{C}]} - r[\mathsf{C}]f_1([\mathsf{C}]) + \mathcal{O}\left(r^2\right) \tag{9}$$

where the first order contribution $f_1([\mathsf{C}])$ is a strictly positive function of $[\mathsf{C}]$. The r-independent part has the coefficients:

$$\begin{aligned}
\alpha &= a\bar{b}df + ab\bar{d}h + abfh[\mathsf{A}][\mathsf{B}] \\
\beta &= a\bar{b}\bar{d} + \bar{b}\bar{d}\bar{f} + \bar{b}\bar{d}\bar{h} + \bar{b}f\bar{h}[\mathsf{A}][\mathsf{B}] \\
\gamma &= adfh \\
\delta &= (a + \bar{f})\bar{d}h
\end{aligned} \tag{10}$$

For small values of $c \ll 1$ we have $c \approx p_1[\mathsf{C}]/q_0$ and hence

$$\dot{c} = \frac{q_0^2\alpha}{p_1^2\beta}[\mathsf{A}][\mathsf{B}]c^2 - rc, \tag{11}$$

i.e., second order autocatalytic, i.e., *hyperbolic growth.*

For large values of c, and assuming that we can neglect the r-dependent parts as argued below in subsection 3.4, we have $c \approx (p_4/q_1)[\mathsf{C}]^3$ and hence

$$\dot{c} = \sqrt[3]{\frac{q_1^2}{p_4^2}}\frac{\gamma}{\delta}[\mathsf{A}][\mathsf{B}]c^{2/3} - rc\,. \tag{12}$$

There we observe *parabolic growth.*

In a high flux rate regime, in which $q_2[\mathsf{C}]^2$ cannot be neglected, we have $c \approx (p_4/q_2)[\mathsf{C}]^2$ for small values of c, and hence

$$\dot{c} = \frac{q_2\gamma}{p_4\delta}c - rc \tag{13}$$

corresponding to *exponential growth.*

This brief discussion extends von Kiedrowski's minimal replicator theory [47] from a replication step proceeding via ABC to a mechanism involving ABCC as the crucial intermediate. In addition to parabolic and exponential growth we observe a regime of hyperbolic growth that is characteristic for higher order autocatalysis [6].

3. Modeling Competition

3.1. ***Reaction Mechanism and Kinetic Equations.*** We assume that there are two different classes of short template molecules each of which occur in different versions (sequence variants) $\mathsf{A_i}$, $i = 1, \ldots, M$ and $\mathsf{B_j}$, $j = 1, \ldots, N$. From these building blocks all possible combinations of templates $\mathsf{C_{ij}}$ can be produced. This scenario parallels the peptide-based experiments described in [23, 51].

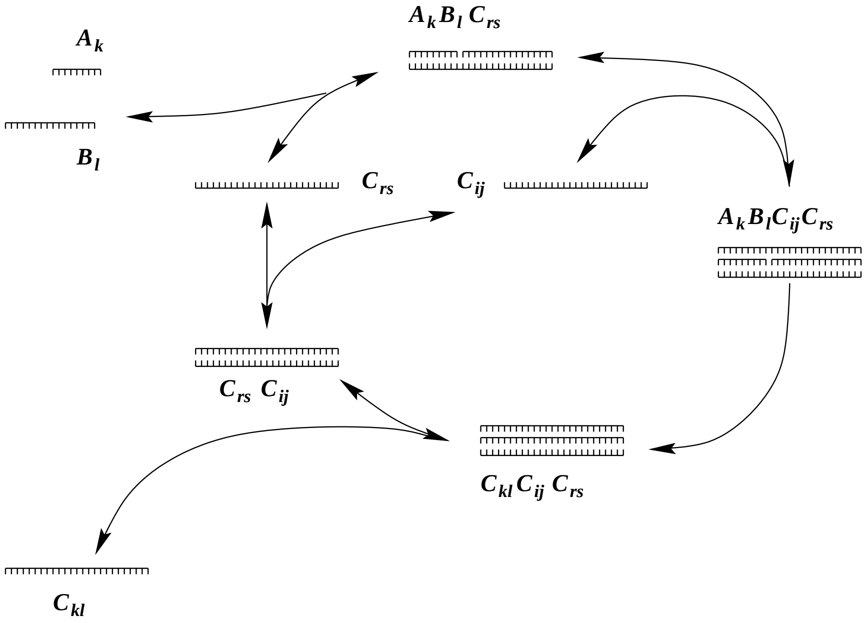

Figure 1. Schematic representation of the full 2nd order ligation-replication mechanism.

The complete reaction scheme of a second order ligation-replication with multiple templates is shown in Figure 1. The corresponding chemical reaction equations are

$$
\begin{aligned}
[\mathsf{A_i}] + [\mathsf{B_j}] + [\mathsf{C_{kl}}] &\underset{\bar{b}_{ijkl}}{\overset{b_{ijkl}}{\rightleftharpoons}} [\mathsf{A_iB_jC_{kl}}] \\
[\mathsf{A_i}] + [\mathsf{B_j}] + [\mathsf{C_{kl}C_{rs}}] &\underset{\bar{f}_{ijklrs}}{\overset{f_{ijklrs}}{\rightleftharpoons}} [\mathsf{A_iB_jC_{kl}C_{rs}}] \\
[\mathsf{A_iB_jC_{kl}}] + [\mathsf{C_{rs}}] &\underset{\bar{h}_{ijklrs}}{\overset{h_{ijklrs}}{\rightleftharpoons}} [\mathsf{A_iB_jC_{kl}C_{rs}}] \\
[\mathsf{C_{kl}}] + [\mathsf{C_{rs}}] &\underset{\bar{d}_{klrs}}{\overset{d_{klrs}}{\rightleftharpoons}} [\mathsf{C_{kl}C_{rs}}] \\
[\mathsf{C_{ij}}] + [\mathsf{C_{kl}C_{rs}}] &\underset{\bar{g}_{ijklrs}}{\overset{g_{ijklrs}}{\rightleftharpoons}} [\mathsf{C_{ij}C_{kl}C_{rs}}] \\
[\mathsf{A_iB_jC_{kl}C_{rs}}] &\xrightarrow{a_{ijklrs}} [\mathsf{C_{ij}C_{kl}C_{rs}}]
\end{aligned}
\tag{14}
$$

for all $i, k, r \in \{1, \dots, M\}$ and $j, l, s \in \{1, \dots, N\}$.

We assume here that $\mathsf{A_iB_jC_{kl}C_{rs}} = \mathsf{A_iB_jC_{rs}C_{kl}}$, $\mathsf{C_{kl}C_{ij}} = \mathsf{C_{ij}C_{kl}}$, $\mathsf{C_{ij}C_{kl}C_{rs}} = \mathsf{C_{kl}C_{ij}C_{rs}}$, etc. It will be convenient therefore to introduce a lexicographic order for the index pairs, defined by $(i, j) < (k, l)$ if $i < k$ or $i = k$ and

$j < l$. In the following we will use the abbreviation $ij < kl$. A sum over all species of the type $\mathsf{C_{kl}C_{ij}}$ can now simply be written as $\sum_{ij<kl}$. The kinetic equations for the intermediates, in the cases where the index pairs of the template species, kl and rs, are pairwisely distinct, read

$$\begin{aligned}
\frac{\mathrm{d}[\mathsf{A_iB_jC_{kl}}]}{\mathrm{d}t} &= b_{ijkl}[A_i][B_j][C_{kl}] - \bar{b}_{ijkl}[\mathsf{A_iB_jC_{kl}}] + \sum_{r,s}\bar{h}_{ijklrs}[\mathsf{A_iB_jC_{kl}C_{rs}}] \\
&\quad - \sum_{r,s} h_{ijklrs}[\mathsf{A_iB_jC_{kl}}][\mathsf{C_{rs}}] - r[\mathsf{A_iB_jC_{kl}}] = 0 \\
\frac{\mathrm{d}[\mathsf{A_iB_jC_{kl}C_{rs}}]}{\mathrm{d}t} &= f_{ijklrs}[\mathsf{A_i}][\mathsf{B_j}][\mathsf{C_{kl}C_{rs}}] - \bar{f}_{ijklrs}[\mathsf{A_iB_jC_{kl}C_{rs}}] \\
&\quad + h_{ijklrs}[\mathsf{A_iB_jC_{kl}}][\mathsf{C_{rs}}] + h_{ijrskl}[\mathsf{A_iB_jC_{rs}}][\mathsf{C_{kl}}] \\
&\quad - \bar{h}_{ijklrs}[\mathsf{A_iB_jC_{kl}C_{rs}}] - \bar{h}_{ijrskl}[\mathsf{A_iB_jC_{kl}C_{rs}}] \\
&\quad - a_{ijklrs}[\mathsf{A_iB_jC_{kl}C_{rs}}] - r[\mathsf{A_iB_jC_{kl}C_{rs}}] = 0 \\
\frac{[\mathsf{C_{kl}C_{rs}}]}{\mathrm{d}t} &= \sum_{i,j}\bar{f}_{ijklrs}[\mathsf{A_iB_jC_{kl}C_{rs}}] - \sum_{i,j} f_{ijklrs}[\mathsf{A_i}][\mathsf{B_j}][\mathsf{C_{kl}C_{rs}}] \\
&\quad + d_{klrs}[\mathsf{C_{kl}}][\mathsf{C_{rs}}] - \bar{d}_{klrs}[\mathsf{C_{kl}C_{rs}}] + \sum_{i,j}\bar{g}_{ijklrs}[\mathsf{C_{ij}C_{kl}C_{rs}}] \\
&\quad - \sum_{i,j} g_{ijklrs}[\mathsf{C_{ij}}][\mathsf{C_{kl}C_{rs}}] - r[\mathsf{C_{kl}C_{rs}}] = 0 \\
\frac{[\mathsf{C_{ij}C_{kl}C_{rs}}]}{\mathrm{d}t} &= g_{ijklrs}[\mathsf{C_{ij}}][\mathsf{C_{kl}C_{rs}}] + g_{klijrs}[\mathsf{C_{kl}}][\mathsf{C_{ij}C_{rs}}] + g_{rsijkl}[\mathsf{C_{rs}}][\mathsf{C_{ij}C_{kl}}] \\
&\quad - \bar{g}_{ijklrs}[\mathsf{C_{ij}C_{kl}C_{rs}}] - \bar{g}_{klijrs}[\mathsf{C_{ij}C_{kl}C_{rs}}] - \bar{g}_{rsijkl}[\mathsf{C_{ij}C_{kl}C_{rs}}] \\
&\quad + a_{ijklrs}[\mathsf{A_iB_jC_{kl}C_{rs}}] + a_{klijrs}[\mathsf{A_kB_lC_{ij}C_{rs}}] \\
&\quad + a_{rsijkl}[\mathsf{A_rB_sC_{ij}C_{kl}}] - r[\mathsf{C_{ij}C_{kl}C_{rs}}] = 0
\end{aligned} \tag{15}$$

The analogous expression for species consisting of two or three equal template species are slighly different. The expression for $\mathrm{d}[\mathsf{C_{kl}C_{kl}}]/\mathrm{d}t$ does not change. However we have

$$\begin{aligned}
\frac{\mathrm{d}[\mathsf{A_iB_jC_{kl}C_{kl}}]}{\mathrm{d}t} &= f_{ijklkl}[\mathsf{A_i}][\mathsf{B_j}][\mathsf{C_{kl}C_{kl}}] - \bar{f}_{ijklkl}[\mathsf{A_iB_jC_{kl}C_{kl}}] \\
&\quad + h_{ijklkl}[\mathsf{A_iB_jC_{kl}}][\mathsf{C_{kl}}] - \bar{h}_{ijklkl}[\mathsf{A_iB_jC_{kl}C_{kl}}] \\
&\quad - a_{ijklkl}[\mathsf{A_iB_jC_{kl}C_{kl}}] - r[\mathsf{A_iB_jC_{kl}C_{kl}}] = 0
\end{aligned} \tag{16}$$

and similarly

$$\begin{aligned}
\frac{\mathrm{d}[\mathsf{C_{kl}C_{kl}C_{rs}}]}{\mathrm{d}t} &= g_{klklrs}[\mathsf{C_{kl}}][\mathsf{C_{kl}C_{rs}}] + g_{rsklkl}[\mathsf{C_{rs}}][\mathsf{C_{kl}C_{kl}}] \\
&\quad - \bar{g}_{klklrs}[\mathsf{C_{kl}C_{kl}C_{rs}}] - \bar{g}_{rsklkl}[\mathsf{C_{kl}C_{kl}C_{rs}}] \\
&\quad + a_{klklrs}[\mathsf{A_k B_l C_{kl} C_{rs}}] + a_{rsklkl}[\mathsf{A_r B_s C_{kl} C_{kl}}] \\
&\quad - r[\mathsf{C_{kl}C_{kl}C_{rs}}] = 0 \\
\frac{\mathrm{d}[\mathsf{C_{kl}C_{kl}C_{kl}}]}{\mathrm{d}t} &= g_{klklkl}[\mathsf{C_{kl}}][\mathsf{C_{kl}C_{kl}}] - \bar{g}_{klklkl}[\mathsf{C_{kl}C_{kl}C_{kl}}] \\
&\quad + a_{klklkl}[\mathsf{A_k B_l C_{kl} C_{kl}}] - r[\mathsf{C_{kl}C_{kl}C_{kl}}] = 0
\end{aligned} \tag{17}$$

For the concentration $[\mathsf{C_{kl}}]$ of the free replicating template itself we find

$$\begin{aligned}
\frac{\mathrm{d}[\mathsf{C_{kl}}]}{\mathrm{d}t} = \sum_{i,j} \bar{b}_{ijkl}[\mathsf{A_i B_j C_{kl}}] - \sum_{i,j} b_{ijkl}[\mathsf{A_i}][\mathsf{B_j}][\mathsf{C_{kl}}] + \sum_{i,j,r,s} \bar{h}_{ijrskl}[\mathsf{A_i B_j C_{kl} C_{rs}}] \\
- \sum_{i,j,r,s} h_{ijrskl}[\mathsf{A_i B_j C_{rs}}][\mathsf{C_{kl}}] + \sum_{r,s} \bar{d}_{klrs}[\mathsf{C_{kl}C_{rs}}] - \sum_{r,s} d_{klrs}[\mathsf{C_{kl}}][\mathsf{C_{rs}}] \\
+ \bar{d}_{klkl}[\mathsf{C_{kl}C_{kl}}] - d_{klkl}[\mathsf{C_{kl}}][\mathsf{C_{kl}}] + \sum_{\substack{i,j,r,s \\ ij \le rs}} \bar{g}_{klijrs}[\mathsf{C_{ij}C_{kl}C_{rs}}] \\
- \sum_{\substack{i,j,r,s \\ ij \le rs}} g_{klijrs}[\mathsf{C_{kl}}][\mathsf{C_{ij}C_{rs}}] \,.
\end{aligned} \tag{18}$$

A very tedious computation eventually leads to the fairly simple expression

$$\frac{\mathrm{d}c_{kl}}{\mathrm{d}t} = \frac{\mathrm{d}[\mathsf{C_{kl}}]}{\mathrm{d}t} = \sum_{\substack{i,j,r,s \\ ij \le rs}} a_{klijrs}[\mathsf{A_k B_l C_{ij} C_{rs}}] - r\, c_{kl} \tag{19}$$

where c_{kl} denotes the total concentration of the templates in the system. The problem hence reduces to expressing $[\mathsf{A_k B_l C_{ij} C_{rs}}]$ in terms of the total replicator concentrations c_{kl}. To this end we first have to solve eqns.(15,16,17) for $[\mathsf{C_{kl}}]$, then we have to solve

$$\begin{aligned}
c_{kl} = [\mathsf{C_{kl}}] + \sum_{i,j} [\mathsf{A_i B_j C_{kl}}] + \sum_{ij} [\mathsf{C_{ij} C_{kl}}] + [\mathsf{C_{kl} C_{kl}}] + \sum_{i,j} \sum_{rs \le kl} [\mathsf{A_i B_j C_{rs} C_{kl}}] \\
+ \sum_{i,j} [\mathsf{A_i B_j C_{kl} C_{kl}}] + \sum_{ij \le rs \le kl} [\mathsf{C_{ij} C_{rs} C_{kl}}] + \sum_{ij \le kl} [\mathsf{C_{ij} C_{kl} C_{kl}}] + [\mathsf{C_{kl} C_{kl} C_{kl}}] \,.
\end{aligned} \tag{20}$$

Not surprisingly, explicit solutions of these systems of algebraic equations cannot be obtained in general. In most cases, hence, we will have to resort to qualitative results.

3.2. *Transformation to Replicator Form.* Denoting the total replicator concentration by $c_0 = \sum_j c_j$ we define the relative concentrations

$$x_{kl} = c_{kl}/c_0 \qquad \text{and} \qquad y_{kl} = [\mathsf{C_{kl}}]/c_0\,. \tag{21}$$

As an immediate consequence of the discussion at the end of section 3.1 we have

$$a_{klijrs}[\mathsf{A_k B_l C_{ij} C_{rs}}] = G_{kl,ij,rs}(\mathbf{y}; c_0) = F_{kl,ij,rs}(\mathbf{x}; c_0)\,. \tag{22}$$

Next we compute

$$\begin{aligned}\dot{x}_{kl} = \frac{\mathrm{d}x_{kl}}{\mathrm{d}t} &= \frac{1}{c_0}\left(\frac{\mathrm{d}c_{kl}}{\mathrm{d}t} - \frac{c_{kl}}{c_0}\sum_{p,q}\frac{\mathrm{d}c_{pq}}{\mathrm{d}t}\right) = \frac{1}{c_0}\sum_{ij\le rs} F_{kl,ij,rs}(\mathbf{x}; c_0) - r\frac{c_{kl}}{c_0} \\ &- x_{kl}\sum_{\substack{pq,ij,rs\\ ij\le rs}}\frac{1}{c_0}F_{pq,ij,rs}(\mathbf{x}; c_0) + x_{kl}\, r\sum_{pq}\frac{c_{pq}}{c_0}\,.\end{aligned} \tag{23}$$

The two r-dependent terms cancel. Next we split the remaining sums into two terms

$$\widehat{F}^{\mathrm{repl}}_{kl}(\mathbf{x}; c_0) = \sum_{ij} F_{kl,kl,ij}(\mathbf{x}; c_0) \qquad \widehat{F}^{\mathrm{mut}}_{kl,ij}(\mathbf{x}; c_0) = \sum_{\substack{rs:\, ij\le rs\\ ij\ne kl,\, rs\ne kl}} F_{kl,ij,rs}(\mathbf{x}; c_0)\,. \tag{24}$$

The term $\widehat{F}^{\mathrm{repl}}_{kl}(\mathbf{x}; c_0)$ describes the total production rate of $\mathsf{C_{kl}}$ through intermediates that involve $\mathsf{C_{kl}}$ as a template. We call these the *replication* pathways. The term $\widehat{F}^{\mathrm{mut}}_{kl,ij}(\mathbf{x}; c_0)$, on the other hand, summarizes the production of $\mathsf{C_{kl}}$ along pathways that use another template $\mathsf{C_{ij}}$. We call these the *mutation* or miscopying pathways. (Since there is no actual distinction between the role of $\mathsf{C_{ij}}$ and $\mathsf{C_{rs}}$ we use the lexicographic ordering of the indices to avoid double counting. Note that this is a mere book-keeping device.)

Define the *replication rate function*

$$\varphi_{kl}(\mathbf{x}) := \frac{1}{c_{kl}}\sum_{i,j} a_{klklij}[\mathsf{A_k B_l C_{kl} C_{ij}}] = \frac{1}{x_{kl}}\widehat{F}^{\mathrm{repl}}_{kl}(\mathbf{x}, c_0) \tag{25}$$

and the *mutation field*

$$\mathcal{M}_{kl}(\mathbf{x}) = \frac{1}{c_0}\sum_{\substack{i,j\\ ij\ne kl}}\left(\widehat{F}^{\mathrm{mut}}_{kl,ij}(\mathbf{x}; c_0) - x_{kl}\sum_{\substack{p,q\\ pq\ne ij}}\widehat{F}^{\mathrm{mut}}_{ij,kl}(\mathbf{x}; c_0)\right) \tag{26}$$

It is not hard to check that $\mathcal{M}$ is a mutation vector field in the sense of the perturbation theory developed in [42]. A short computation verifies that equ.(23) can be written in the form

$$\dot{x}_{kl} = x_{kl}\left[\varphi_{kl}(\mathbf{x}) - x_{kl}\sum_{pq} x_{pq}\varphi_{pq}(\mathbf{x})\right] + \mathcal{M}_{kl}(\mathbf{x}) \tag{27}$$

Apart from the mutation term $\mathcal{M}_{kl}(\mathbf{x})$, Equ.(27) is the general form of a *replicator equation* as introduced in [32] as the general model of competition under *constant organization.*

3.3. *Some Technical Details.* In the subsection 3.2 we have tacitly assumed that $G_{kl,ij,rs}(\mathbf{y}; c_0)$ and $F_{kl,ij,rs}(\mathbf{x}; c_0)$ are well defined.

The quasi-steady state approximation (QSSA) equ.(15,16,17) is valid only in the range of parameters, for concentrations $[\mathsf{C_{kl}}]$ and flux rates r in which the solution for $[\mathsf{A_iB_jC_{kl}}]$ etc. is stable; for a thorough discussion of the QSSA see [2, 37]. Note that the diagonal entries of the Jacobian of equ.(15) are always negative, $\partial\frac{\mathrm{d}[\mathsf{A_iB_jC_{kl}}]}{\mathrm{d}t}/\partial[\mathsf{A_iB_jC_{kl}}] = -(\bar{b}_{ijkl} + r)$, etc. Thus there is a parameter regime in which this matrix is diagonally dominated and hence stable. Within this parameter domain there is a unique solution of equ.(15) that depends smoothly on $[\mathsf{C_{kl}}]$ for all kl in a concentration domain that includes the origin. Thus, within the range of QSSA-validity, $G_{kl,ij,rs}(\mathbf{y}; c_0)$ is a well defined *smooth* function.

For small enough c_0, a rather tedious computations shows that $[\mathsf{A_iB_jC_{kl}}] = \mathcal{O}([\mathsf{C_{kl}}])$, $[\mathsf{A_iB_jC_{kl}C_{rs}}] = \mathcal{O}([\mathsf{C_{kl}}][\mathsf{C_{rs}}])$, $[\mathsf{C_{kl}C_{rs}}] = \mathcal{O}([\mathsf{C_{kl}}][\mathsf{C_{rs}}])$, $[\mathsf{C_{ij}C_{kl}C_{rs}}] = \mathcal{O}([\mathsf{C_{ij}}][\mathsf{C_{rs}}]) + \mathcal{O}([\mathsf{C_{kl}}][\mathsf{C_{rs}}]) + \mathcal{O}([\mathsf{C_{ij}}][\mathsf{C_{rs}}])$. Of course the coefficients must be positive, since we are dealing with non-vanishing concentrations. Thus $\partial c_{kl}/\partial[\mathsf{C_{kl}}] = \mathcal{O}(1) + \mathcal{O}([\mathsf{C}])$, where $\mathcal{O}([\mathsf{C}])$ stands for a term proportional to the concentrations of the free templates, i.e., $\mathcal{O}([\mathsf{C}]) \leq \mathcal{O}(c_0)$. On the other hand, the off-diagonal terms $\partial c_{kl}/\partial[\mathsf{C_{rs}}] = \mathcal{O}([\mathsf{C}]) \leq \mathcal{O}(c_0)$, hence this matrix, the Jacobian $\mathbf{J}$ of the map $([\mathsf{C_{kl}}]) \mapsto (c_{kl})$ is diagonally dominated and hence positive definite. At least for small enough c_0 this guarantees existence and uniqueness of a smooth function $\mathbf{y}(\mathbf{x}, c_0)$; this function can be extended at least into the entire regime in which $\mathbf{J}$ stays regular. We strongly suspect that for every choice of template concentrations (c_{kl}) there is a unique solution $([\mathsf{C_{kl}}])$, although we have not been able to proof this for the general case.

Finally, definition of the relative growth rate $\varphi_{kl}(\mathbf{x})$ in equ.(25) suggests that φ_{kl} might behave badly for small x_{kl}. If this were indeed the case, the transformation of eqns(15,16,17,19) to replicator form would have been a rather useless exercise. Fortunately, we may argue as follows: $[\mathsf{A_kB_lC_{kl}C_{rs}}] \leq c_{kl}$ and hence $0 < [\mathsf{A_kB_lC_{kl}C_{rs}}]/c_{kl} < 1$ for all c_{kl}. Thus φ_{kl} is indeed well defined. A rather lengthy calculation shows furthermore that $[\mathsf{A_kB_lC_{kl}C_{rs}}] \propto$

$[\mathsf{C_{kl}}]$ and $[\mathsf{A_k B_l C_{kl} C_{kl}}] \propto [\mathsf{C_{kl}}]^2$ for $[\mathsf{C_{kl}}] \ll 1$, hence φ_{kl} is indeed a smooth function on the concentration simplex.

3.4. ***An Approximation.*** In [14] it is shown that the dynamics of a reaction network in a CSTR is essentially equivalent with replicator dynamics under constant organization, i.e., with constant c_0 provided the CSTR's flux rate r is small enough. In particular, the discrepancy is a regular perturbation in the sense of [16], which does not affect the behavior of structurally stable phase portraits. Assuming small r will also allow us to neglect the r-dependent terms in eqns.(15,16,17), making their solution much more tractable. In the remainder of this contribution we will make this approximation.

4. Independent Replication

As a first example consider a collection of independently replicating template species $\mathsf{C_{kl}}$ that do not form "mixed" intermediates such as $\mathsf{A_i B_j C_{kl}}$ or $\mathsf{A_i B_j C_{kl} C_{rs}}$ with $ij \neq kl$, $rs \neq kl$. Assuming that these molecules simply compete for space in the flow reactor, we see that each species replicates according to the kinetics of a single template as detailed in section 2. It will be convenient to replace the double index kl by k. We have explicitly

$$\dot{c}_k = [\mathsf{C_k}]^2 \frac{\alpha_k + \gamma_k [\mathsf{C_k}]}{\beta_k + \delta_k [\mathsf{C_k}]} - r c_k \tag{28}$$

where c_k and $[\mathsf{C_k}]$ are related via a rational equation analogous to equ.(8). In replicator form we find

$$\dot{x}_k = x_k \left[\varphi_k(x_k) - \sum_j x_j \varphi_j(x_j) \right] \tag{29}$$

where

$$\begin{aligned} \varphi_k(x_k) &\sim c_0 \frac{q_{k;0}^2 \alpha_k}{p_{k;1}^2 \beta_k} [\mathsf{A_k}][\mathsf{B_k}] x_k & \qquad c_0 \to 0 \\ \varphi_k(x_k) &\sim \sqrt[3]{\frac{q_{k;1}^2}{c_0 p_{k;4}^2} \frac{\gamma_k}{\delta_k}} [\mathsf{A_k}][\mathsf{B_k}] x_k^{-1/3} & \qquad c_0 \to \infty \end{aligned} \tag{30}$$

As a consequence, we recover two well studied models as limiting cases: In dilute solution, $c_0 \to 0$, we obtain Schlögel's model of competing autocatalysts [31], in saturated solution, $c_0 \to \infty$, we recover Szathmáry's model of parabolic growth [43].

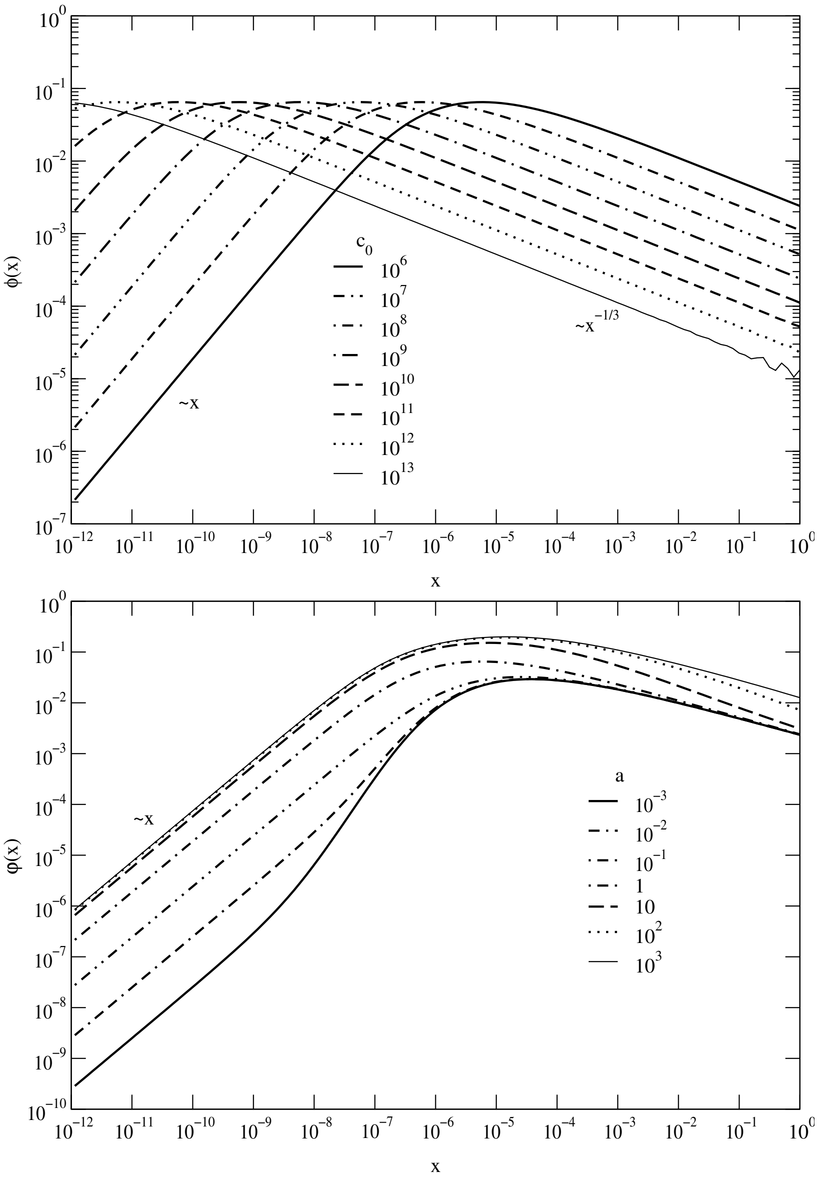

Figure 2. Interaction function $\varphi(x)$ for the example system equ.(31) with $a = 1$ and various values of c_0 (upper panel) and $c_0 = 10^6$ and various values of a (lower panel), respectively. The logarithmic plot emphasizes the two asymptotic regimes $\varphi(x) \sim x$ for small x and $\varphi(x) \sim x^{-1/3}$ for large x.

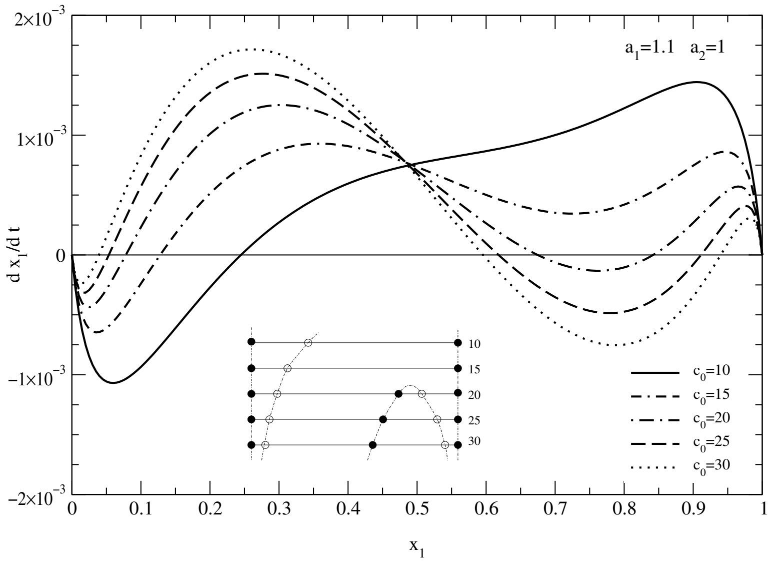

Figure 3. Bifurcation in a two-species model with independent replicators. The transition from a Schlögel-type phase portrait with two stable corner equilibria and an unstable interior equilibrium to a Szathmáry-like behavior exhibiting a stable interior equilibrium and unstable equilibria very close to both corners proceeds via a saddle-node bifurcation in the interior of the simplex.

Example. Suppose we set all constants except the replication rate a_k to 1. Then:

$$\begin{array}{llll} \alpha_k = 3a_k & \beta_k = a_k + 3 & \gamma_k = 1 \quad \delta_k = 2 \\ p_{4;k} = 3(a_k + 1) & p_{3;k} = 11a_k + 10 & p_{2;k} = 14(a_k + 1) \\ p_{1;k} = 2(a_k + 3) & q_{1;k} = a_k + 1 & q_{0;k} = a_k + 3 \end{array} \tag{31}$$

The replicator equation (29) is a Shashahani gradient [17], hence all ω-limits are fixed points. All corners of the simplex, i.e., the points $\mathbf{e}_k$ with $x_k = 1$ and $x_i = 0$, $i \neq k$ are fixed points. Their stability is determined by the "transversal eigenvalues"

$$\lambda_i(\mathbf{e}_k) = -\varphi_i([\mathbf{e}_k]_i) = -\varphi_i(0) < 0 \tag{32}$$

The fixed points on the corner are therefore always asymptotically stable. Their basin of attraction, which is essentially determined by the area around $\mathbf{e}_k$ in which the transversal eigenvalues are still negative, may be tiny, as Figure 2 shows.

Figure 3 shows a typical bifurcation diagram for the two-species case. In contrast to the 2nd order replicator equations we observe generic saddle-node bifurcations. The transition from the bi-stable case of initial-condition dependent selection to a Szathmáry-like behavior is initiated by a saddle-node bifurcation that creates a source-sink pair in the interior of S_2. As

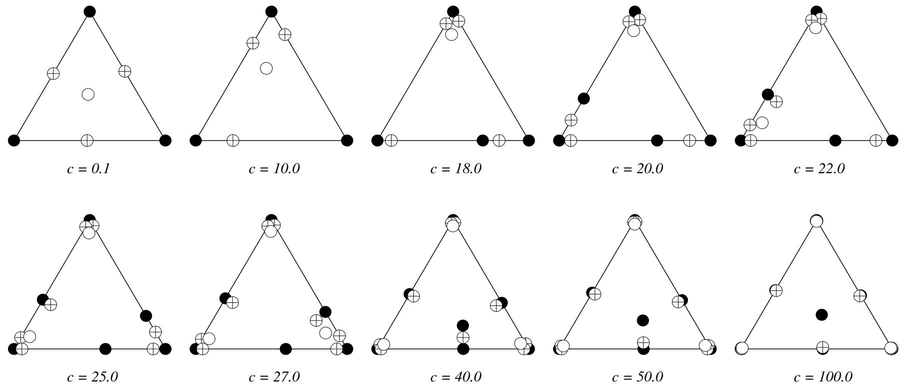

Figure 4. The transition from competition to cooperation in a system containing three independently replicating species proceeds via a complex series of saddle-node bifurcations. Parameters: $a_1 = 1.1$, $a_2 = 1.2$, $a_3 = 1.0$. $x_1 = 1$ is the lower left corner, $x_3 = 1$ the top of the triangle.

c_0 increases, the unstable equilibria move closer and closer to the corners, effectively "shielding" the effect of the stable rest points in both corners.

Figure 4 shows a typical bifurcation diagram for the three-species case. For small c_0 we have the well-known phase portrait of Schlögel's model. As c_0 increases the saddles on the edge move towards the corners and saddle-node bifurcations create saddle-sink pairs on the edges. Around $c_0 = 10$ the phase portrait resembles the "survival of the fittest" like behavior characteristic for exponential growth. There is a dominating basin of attraction for $\mathbf{e}_2$, corresponding to selection of the species with the largest a-value, $a_2 = 1.2$. Between $c_0 = 20$ and $c_0 = 22$ the first saddle-node bifurcation in the interior of S_3 occurs, creating a source-saddle pair. An analogous bifurcation occurs between $c_0 = 25$ and $c_0 = 27$. The last bifurcation creates a saddle-sink pair. As c_0 increases further, this sink moves towards the center of the simplex. The phase portrait approaches more and more the familiar picture of the parabolic growth models with its globally stable interior equilibrium.

5. Correct Instruction

5.1. ***Simplifications.*** The templates in the general reaction mechanism (14) are promiscuous in the sense that, in principle, the molecular species C_{ij} can be composed from A_i and B_j using other species C_{ij}, $ij \neq kl$, as template and catalyst. Hence the mutation term $\mathcal{M}$ does not vanish. In this sense, equ.(14) therefore corresponds to a selection-mutation system as described e.g. in [42] or a general catalytic network [40], not to a network of replicators.

Template instructed replication in its simplest form is dependent upon a

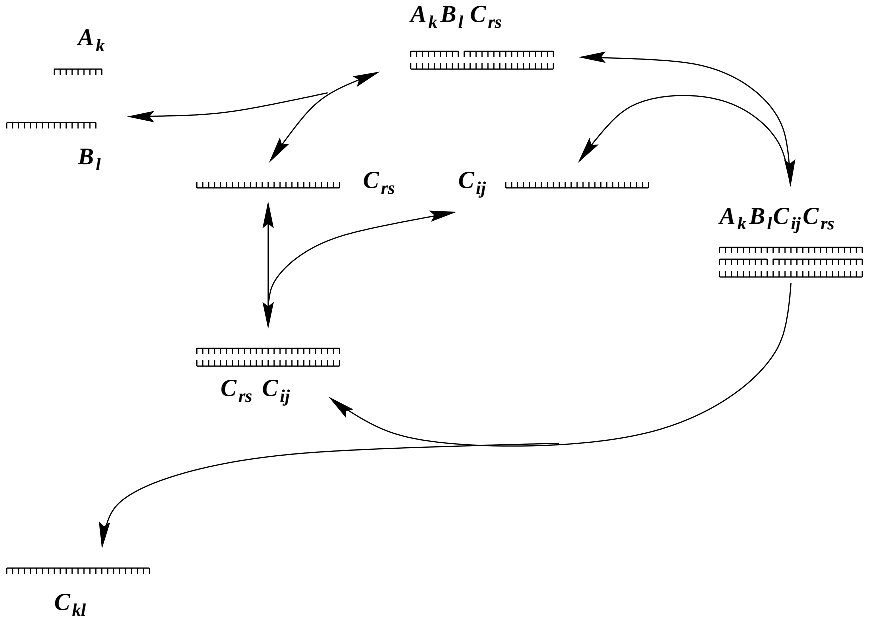

Figure 5. A simplified reaction scheme assuming that the ternary template complexes $\mathsf{C_{ij}C_{kl}C_{rs}}$ are so unstable that they immediately disintegrate.

specific recognition between template molecules $\mathsf{C_{kl}}$ and the building material $\mathsf{A_i}$ and $\mathsf{B_j}$. This requirement can be built into the present model by restricting the types of intermediates that occur with a non-zero concentration, i.e., by restricting the reaction pathways. We remark that the mutation term $\mathcal{M}$ vanishes already when $a_{klijrs} = 0$ for $kl \neq ij$ and $kl \neq rs$. It does seem realistic however, to assume that the complexes $\mathsf{A_kB_lC_{ij}}$ are unstable unless $ij = kl$, i.e., unless there is proper molecular recognition. As a consequence then, only intermediates of the form $\mathsf{A_kB_lC_{kl}}$, $\mathsf{A_kB_lC_{kl}C_{kl}}$, and $\mathsf{A_kB_lC_{kl}C_{rs}}$, $rs \neq kl$, can be formed.

The resulting set of algebraic equations for the intermediate species is still too complicated to allow for a detailed analytic treatment. Hence we resort to the simplification outlined in Figure 5. Here we assume that the ternary complexes $\mathsf{C_{kl}C_{kl}C_{rs}}$ decompose very fast, so that the irreversible reaction step becomes

$$\mathsf{A_kB_lC_{kl}C_{rs}} \xrightarrow{a_{klklrs}} \mathsf{C_{kl}} + \mathsf{C_{rs}C_{kl}} \tag{33}$$

where the newly formed template molecule $\mathsf{C_{kl}}$ is released in its free form.

It is straight forward to derive the kinetic equations for the simplified reaction mechanism under the assumption of correct instruction. Explicitly we have

$$
\begin{aligned}
\frac{\mathrm{d}[\mathsf{A_k B_l C_{kl}}]}{\mathrm{d}t} &= b_{klkl}[\mathsf{A_k}][\mathsf{B_l}][\mathsf{C_{kl}}] \\
&\quad - \left(\bar{b}_{klkl} + h_{klklkl}[\mathsf{C_{kl}}] + \sum_{\substack{i,j\\ ij\neq kl}} h_{klklij}[\mathsf{C_{ij}}]\right)[\mathsf{A_k B_l C_{kl}}] \\
&\quad + \sum_{\substack{i,j\\ ij\neq kl}} \bar{h}_{klklij}[\mathsf{A_k B_l C_{kl} C_{ij}}] + \bar{h}_{klklkl}[A_k B_l C_{kl} C_{kl}] = 0 \\
\frac{\mathrm{d}[\mathsf{A_k B_l C_{kl} C_{kl}}]}{\mathrm{d}t} &= h_{klklkl}[\mathsf{A_k B_l C_{kl}}][\mathsf{C_{kl}}] - (\bar{h}_{klklkl} + a_{klklkl})[\mathsf{A_k B_l C_{kl} C_{kl}}] = 0 \\
\frac{\mathrm{d}[\mathsf{A_k B_l C_{kl} C_{rs}}]}{\mathrm{d}t} &= h_{klklrs}[A_k B_l C_{kl}][C_{rs}] \\
&\quad - (\bar{h}_{klrskl} + \bar{h}_{klklrs} + a_{klklrs})[\mathsf{A_k B_l C_{kl} C_{rs}}] = 0 \\
\frac{\mathrm{d}[\mathsf{C_{kl} C_{rs}}]}{\mathrm{d}t} &= a_{klklrs}[\mathsf{A_k B_l C_{kl} C_{rs}}] + a_{rsrskl}[\mathsf{A_r B_s C_{rs} C_{kl}}] \\
&\quad + d_{klrs}[\mathsf{C_{kl}}][\mathsf{C_{rs}}] - \bar{d}_{klrs}[\mathsf{C_{kl} C_{rs}}] = 0 \\
\frac{\mathrm{d}[\mathsf{C_{kl} C_{kl}}]}{\mathrm{d}t} &= a_{klklkl}[\mathsf{A_k B_l C_{kl} C_{kl}}] + d_{klkl}[\mathsf{C_{kl}}][\mathsf{C_{kl}}] - \bar{d}_{klkl}[\mathsf{C_{kl} C_{kl}}] = 0
\end{aligned}
\tag{34}
$$

In this simplified system it is possible to solve the steady state equations for the intermediates explicitly. We find

$$
\begin{aligned}
\frac{\mathrm{d}c_{kl}}{\mathrm{d}t} &= \sum_{r,s} a_{klklrs}[\mathsf{A_k B_l C_{kl} C_{rs}}] - r c_{kl} \\
[\mathsf{A_k B_l C_{kl} C_{rs}}] &= \frac{b^*_{klkl}}{1 + \sum_{i,j} h^*_{klij}[\mathsf{C_{ij}}]}[\mathsf{A_k}][\mathsf{B_l}][\mathsf{C_{kl}}][\mathsf{C_{rs}}]
\end{aligned}
\tag{35}
$$

with the aggregate reaction constants

$$
\begin{array}{lll}
\tilde{h}_{klij} = h_{klklij}\left(1 - \frac{\bar{h}_{klklij}}{\bar{h}_{klklij} + \bar{h}_{klijkl} + a_{klklij}}\right) & kl \neq ij & h^*_{klij} = \frac{\tilde{h}_{klij}}{\bar{b}_{klkl}} \\
\tilde{h}_{klkl} = h_{klklkl}\left(1 - \frac{\bar{h}_{klklkl}}{\bar{h}_{klklkl} + a_{klklkl}}\right) & & b^*_{klkl} = \frac{b_{klkl}}{\bar{b}_{klkl}}
\end{array}
\tag{36}
$$

It will be convenient to rewrite this as

$$
\dot{c}_{kl} = [\mathsf{A_k}][\mathsf{B_l}][\mathsf{C_{kl}}]\frac{\sum_{r,s} a^*_{klrs}[C_{rs}]}{1 + \sum_{r,s} h^*_{klrs}[C_{rs}]} - c_{kl} \tag{37}
$$

where

$$
a^*_{klrs} = \frac{a_{klklrs} h_{klklrs} b^*_{klkl}}{\bar{h}_{klrskl} + \bar{h}_{klklrs} + a_{klklrs}} \qquad a^*_{klkl} = \frac{a_{klklkl} h_{klklkl} b^*_{klkl}}{\bar{h}_{klklkl} + a_{klklkl}} \tag{38}
$$

5.2. *Unstable Heteroduplices.* As a limiting case we assume that hetero-duplices are (relatively) unstable, i.e.,

$$h_{klklrs}/\bar{h}_{klklrs} \ll h_{klklkl}/\bar{h}_{klklkl} \qquad \text{and} \qquad d_{klrs}/\bar{d}_{klrs} \ll d_{klkl}/\bar{d}_{klkl}\,, \tag{39}$$

we have $h^*_{klrs} \ll h^*_{klkl}$ and $a^*_{klrs} \ll a^*_{klkl}$. This implies

$$\sum_{i,j} h^*_{klij}[C_{ij}] \approx h^*_{klkl}[C_{kl}] \tag{40}$$

and the relation between c_{kl} and $[\mathsf{C_{kl}}]$ simplifies to

$$c_{kl} = [\mathsf{C_{kl}}]\left\{1 + \tfrac{b^*_{klkl}[\mathsf{A_k}][\mathsf{B_l}]}{1+h^*_{klkl}[\mathsf{C_{kl}}]} + 2[\mathsf{C_{kl}}]\left(\tfrac{d_{klkl}}{\bar{d}_{klkl}} + \tfrac{[A_k][B_l]a^*_{klkl}}{(1+h^*_{klkl}[\mathsf{C_{kl}}])}\tfrac{\bar{d}_{klkl}+a_{klklkl}}{\bar{d}_{klkl}a_{klklkl}}\right)\right\} \tag{41}$$

This can be rewritten as cubic equation

$$c_{kl} = [\mathsf{C_{kl}}]\frac{\alpha_{kl} + \beta_{kl}[\mathsf{C_{kl}}] + \delta_{kl}[\mathsf{C_{kl}}]^2}{1+\gamma_{kl}[\mathsf{C_{kl}}]} \tag{42}$$

We obtain the asymptotic solutions

$$\begin{aligned} &[\mathsf{C_{kl}}] \sim \alpha_{kl}^{-1} c_{kl} && c_0 \to 0 \\ &[\mathsf{C_{kl}}] \sim \sqrt{\gamma_{kl}/\delta_{kl}}\, c_{kl}^{1/2} && c_0 \to \infty \end{aligned} \tag{43}$$

and, since (a^*_{klij}) and (h^*_{klij}) are approximately diagonal, we recover a variant of the independent replication model discussed in detail in section 4.

5.3. *A Hypercycle-Like System.* Suppose $\mathsf{C_{k'l'}}$ acts as the only specific catalyst for replicating $\mathsf{C_{kl}}$, i.e., $a_{klklrs} = 0$ unless $rs = k'l'$. A detailed analytical study for this system is untractable since

$$\begin{aligned} c_{kl} = {} & [\mathsf{C_{kl}}] + \sum_{r,s}[\mathsf{C_{kl}C_{rs}}] + [\mathsf{C_{kl}C_{kl}}] + \\ & + [\mathsf{A_kB_lC_{kl}}] + \sum_{r,s}[\mathsf{A_kB_lC_{kl}C_{rs}}] + \sum_{r,s\neq k,l}[\mathsf{A_rB_sC_{rs}C_{kl}}] \\ & + [\mathsf{A_kB_lC_{kl}C_{kl}}] + \sum_{r,s}[\mathsf{C_{kl}C_{rs}C_{rs}}] + 2\sum_{r,s}[\mathsf{C_{kl}C_{kl}C_{rs}}] \end{aligned} \tag{44}$$

cannot be solved explicitly. Hence we restrict ourselves to a simple symmetric special case: We set $a_{klklk'l'} = a$, $a_{klklrs} = 0$ for $rs \neq k'l'$ and all other rate constants are unity. Thus

$$\begin{aligned} a^*_{klk'l'} &= \frac{a}{2+a} \qquad && a^*_{klrs} = 0 \qquad rs \neq k'l' \\ h^*_{klk'l'} &= \frac{a+3}{a+2} \qquad && h^*_{klrs} = 0 \qquad rs \neq k'l' \end{aligned} \tag{45}$$

and hence

$$\dot{c}_{kl} = [\mathsf{C_{kl}}] \frac{a[\mathsf{C_{k'l'}}]}{2 + a + (3 + a)[\mathsf{C_{k'l'}}]} - c_{kl} \tag{46}$$

The relationship of free and total template concentrations reduces to

$$c_{kl} = [\mathsf{C_{kl}}] \left\{ 1 + [\mathsf{C_{kl}}] + \sum_{rs} [\mathsf{C_{rs}}] + \frac{(1 + a)[\mathsf{C_{k'l'}}]}{2 + a + (3 + a)[\mathsf{C_{k'l'}}]} \right\} \tag{47}$$

We have not been able to determine the explicit solution $y_{kl}(\mathbf{x}, c_0) = [\mathsf{C_{kl}}]$ of equ.(47); hence we have to resort to numerical techniques in general. For c_0 small enough we recover $y_{kl}(\mathbf{x}, c_0) \sim x_{kl}/2$.

For the model (46) we obtain a replicator equation (27) with interaction function

$$\varphi_{kl}(\mathbf{x}) = \frac{y_{kl}(\mathbf{x})}{x_{kl}} \frac{a y_{k'l'}(\mathbf{x})}{\frac{2+a}{c_0} + (3 + a) y_{k'l'}(\mathbf{x})} \tag{48}$$

For small enough c_0 equ.(48) reduces to a hypercycle with the following interaction function

$$\varphi_{kl}(\mathbf{x}) \to \frac{1}{2} \frac{a x_{k'l'}}{\frac{2(2+a)}{c_0} + (3 + a) x_{k'l'}} \tag{49}$$

This is of the form $\varphi_{kl}(\mathbf{x}) = x_{k'l'} g(x_{k'l'})$ where $g(x_{k'l'})$ is continuous and bounded away from 0 on a neighborhood of $[0, 1]$. Hence the hypercycle equ.(49) satisfies the requirement of [17, Thm.B] and we may conclude immediately, that this model is permanent, i.e., for each initial condition with $x_{kl}(0) > 0$, $kl \in \{1, \ldots M\} \times \{1, \ldots, N\}$ there is a constant $\delta > 0$ such that $x_{kl}(t) > \delta$ for sufficiently large time t. For more details on permanent coexistence we refer the reader to [18, 19, 36].

For the large concentration regime we may argue as follows: From $0 < y_{kl}(\mathbf{x}, c_0) < x_{kl}$ for any point $\mathbf{x}$ in the interior of the simplex we see that

$$g_{kl}(\mathbf{x}) = \frac{\varphi_{kl}(\mathbf{x})}{x_{k'l'}} = \frac{y_{kl}(\mathbf{x})}{x_{kl}} \frac{y_{k'l'}(\mathbf{x})}{x_{k'l'}} \frac{a}{\frac{2+a}{c_0} + (3 + a) y_{k'l'}(\mathbf{x})} \tag{50}$$

From equ.(47) we see that $[\mathsf{C_{kl}}]/c_{kl} = y_{kl}(\mathbf{x})/x_{kl}$ is bounded away from 0 since the expression in parentheses is finite for any given c_0. Thus the function $g_{kl}(\mathbf{x})$ defined in equ.(50) is always bounded away from 0 and we can apply [17, Thm.B] to show that our generalized hypercycle is indeed permanent.

We have simulated the hypercyclic system (49) for a different number of species n. For $n = 2, 3, 4$ we find a globally stable interior fixed point

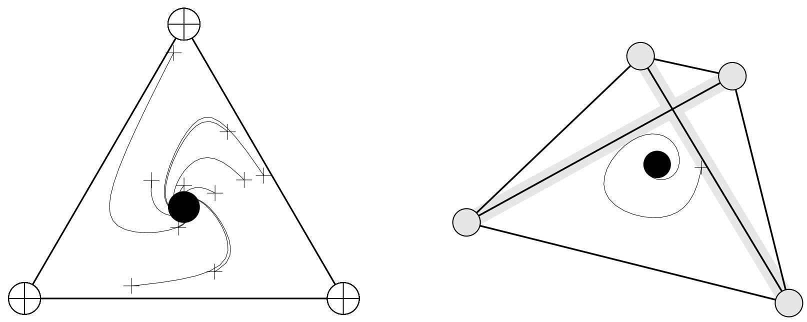

Figure 6. Typical behavior of the hypercyclic model equ.(49) for $n = 3$ and $n = 4$. The behavior does not change qualitatively with the parameters a and c_0. The data for this figure are $a = 1$, $c_0 = 10$. The "non-flowing" edges which consist entirely of degenerate fixed points are indicated by gray bars in the $n = 4$ case.

independent of the parameters a and c_0. We remark that the peculiar inner fixed point of the elementary hypercycle with $n = 4$ which is a marginally stable center in the linear approximation and owes its asymptotic stability to the nonlinear terms only [33] has been replaced here by an ordinary stable fixed point (see also the elementary and the realistic hypercycle in [6]). For $n \geq 4$ the model exhibits *non-flowing* edges $[\mathbf{e}_{ij}, \mathbf{e}_{kl}]$ whenever ij and kl are not adjacent along the hypercycle.

For $n \geq 5$ we observe a Hopf bifurcation at the interior rest point leading from a saddle point and stable limit cycle for small c_0 to a stable interior equilibrium for large concentrations c_0. For $n = 5$ and $a = 1$, the bifurcation point is around $\hat{c}_0 \approx 33$, for instance, for $n = 6$ we find $\hat{c}_0 = 198$.

6. Discussion

A generic class of autocatalytic reaction networks in which replication is based on a multi-step ligation mechanism has been described and analyzed here. Apart from the irreversible ligation step we consider only reversible aggregation steps and thus provide a realistic description of molecular recognition. The reaction mechanism of template induced ligation which is catalyzed by a molecule of the same class is probably the simplest physically feasible way of assembling "molecular ecologies" of strongly interacting replicators, although experimental studies so far have made use of more involved reaction schemes [27, 50].

The kinetic differential equations can be cast in replicator form with highly non-linear response functions. The total concentration of replicating material determines the improtance of product inhibition and hence is a crucial parameter for the system's global behavior.

At low concentration there is a regime of hyperbolic growth in which the models reduce to well-understood second order replicator equations. For large concentration the effects of product inhibition dominate and we observe a regime of parabolic growth.

In contrast to the typical behavior of second order replicator equations we observe here only generic saddle-node and Hopf bifurcations. The highly non-linear form of the interaction functions apparently suppresses the transcritical bifurcations that are ubiquitous in simpler replicator equation models.

Permanent coexistence is a rare phenomenon in the hyperbolic growth regime, where it is restricted to a small parameter range often associated with positive feedback loops such as the hypercycle. Numerical surveys have shown that cooperative networks of this type are exceedingly rare [41] and that they grow extremely slowly only, if at all, through the incorporation of mutants [13].

Product inhibition, and the associated reduction of net growth rates has the effect of favoring coexistence. Even in the case of independently competing species we find global coexistence for large concentrations. Similarly, the unstable interior fixed points of large hypercycles become stable in this regime. It is very likely therefore that coexistence is a much more likely phenomenon in networks of molecular autocatalysts than the estimates based on hyperbolic growth would suggest.

The reaction mechanism (14) already hints at the limitations of studies on the reaction kinetics of large autocatalytic networks. The number of reaction rate constants is excessively large for systems with interesting diversity. Hence the parameters themselves will have to be estimated from a theoretical model. At this point, however, the chemical details play a crucial role, while the global results discussed in this contribution are independent of the particulae chemical implementation.

Acknowledgement

Discussions with Steen Rasmussen are gratefully acknowledged. This work was supported in part by the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, Proj. 13093-GEN.

References

1. D. P. Bartel and J. W. Szostak. Isolation of new ribozymes from a large pool of random sequences. *Science*, 261:1411–1418, 1993.
2. J. A. M. Borghans, R. J. de Boer, and L. A. Segel. Extending the quasi-

steady state approximation by changing variables. *Bull. Math. Biol.*, 58:43–63, 1996.

3. T. R. Cech. RNA splicing: Three themes with variations. *Cell*, 34: 713–716, 1983.
4. M. Eigen. Selforganization of matter and the evolution of macromolecules. *Naturwiss.*, 58:465–523, 1971.
5. M. Eigen, J. S. McCaskill, and P. Schuster. The molecular quasi-species. *Adv. Chem. Phys.*, 75:149–263, 1989.
6. M. Eigen and P. Schuster. *The Hypercycle*. Springer-Verlag, New York, Berlin, 1979.
7. E. H. Ekland and D. P. Bartel. RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature*, 382:373–376, 1996.
8. E. H. Ekland, J. W. Szostak, and D. P. Bartel. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science*, 269:364–370, 1995.
9. S. W. Fox and H. Dose. *Molecular Evolution and the Origin of Life*. Academic Press, New York, 1977.
10. R. F. Gesteland and J. F. Atkins, editors. *The RNA World*. Cold Spring Harbor Laboratory Press, Plainview, NY, 1993.
11. W. Gilbert. The RNA world. *Nature*, 319:618, 1986.
12. C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, and S. Altman. The RNA moiety of Ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35:849–857, 1983.
13. R. Happel and P. F. Stadler. The evolution of diversity in replicator networks. *J. Theor. Biol.*, 195:329–338, 1998.
14. R. Happel and P. F. Stadler. Autocatalytic replication in a CSTR and constant organization. *J. Math. Biol.*, 38:422–434, 1999.
15. R. Hecht, R. Happel, P. Schuster, and P. F. Stadler. Autocatalytic networks with intermediates I: Irreversible reactions. *Math. Biosc.*, 140: 33–74, 1997.
16. M. W. Hirsch and S. Smale. *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press, New York, 1974.
17. J. Hofbauer, P. Schuster, and K. Sigmund. Competition and cooperation in catalytic selfreplication. *J. Math. Biol*, 11:155–168, 1981.
18. J. Hofbauer and K. Sigmund. *Dynamical Systems and the Theory of Evolution*. Cambridge University Press, Cambridge U.K., 1988.
19. V. Hutson and K. Schmitt. Permanence and the dynamics of biological systems. *Math. Biosc.*, 111:1–71, 1992.
20. 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.
21. G. F. Joyce. The rise and fall of the RNA world. *The New Biologist*, 3: 399–407, 1991.

22. D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, and M. R. Ghadiri. A self-replicating peptide. *Nature*, 382:525–528, 1996.
23. D. H. Lee, K. Severin, and M. R. Ghadiri. Autocatalytic networks: the transition from molecular self-replication to ecosystems. *Curr. Opin. Chem. Biol.*, 1:491–496, 1997.
24. D. H. Lee, K. Severin, and M. R. Ghadiri. Dynamic error correction in autocatalytic peptide networks. *Angew. Chem. Int. Ed.*, 37:126–128, 1998.
25. D. H. Lee, K. Severin, Y. Yokobayashi, and M. R. Ghadiri. Emergence of symbiosis in peptide self-replication through a hypercyclic network. *Nature*, 390:591–594, 1997.
26. S. F. Mason. *Chemical evolution. Origin of the elements, molecules, and living systems.* Clarendon Press, Oxford (UK), 1991.
27. J. S. McCaskill. Spatially resolved *in vitro* molecular ecology. *Biophys Chem*, 66:145–158, 1997.
28. L. E. Orgel. RNA catalysis and the origin of life. *J. Theor. Biol.*, 123: 127–149, 1986.
29. L. E. Orgel. Evolution of the genetic apparatus. A review. *Cold Spring Harbor Symposia on Quantitative Biology*, 52:9–16, 1987.
30. L. E. Orgel. Molecular replication. *Nature*, 358:203–209, 1992.
31. F. Schlögel. Chemical reaction models for non-equilibrium phase transitions. *Z. Physik*, 253:147–161, 1972.
32. P. Schuster and K. Sigmund. Replicator dynamics. *J. Theor. Biol.*, 100: 533–538, 1983.
33. P. Schuster, K. Sigmund, and R. Wolff. Dynamical systems under constant organization I: Topological analysis of a family on non-linear differential equations - A model for catalytic hypercycles. *Bull. Math. Biol.*, 40:743–769, 1977.
34. 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, 1998.
35. A. W. Schwartz. Speculation on the RNA precursor problem. *J. Theor. Biol.*, 187:523–527, 1997.
36. H. Sedaghat. General permanence conditions for nonlinear difference equations of higher order. *J. Math. Anal. Appl.*, 213:496–510, 1997.
37. L. A. Segel and M. Slemrod. The quasi-steady state assumption: a case study in perturbation. *SIAM Rev.*, 31:446–477, 1989.
38. D. Sievers and G. von Kiedrowski. Self-replication of complementary nucleotide-based oligomers. *Nature*, 369:221–224, 1994.
39. J. S.Novick, Q. Feng, and T. Tjivikua. Kinetic studies and modeling of a self-replicating system. *J. Amer. Chem. Soc.*, 113(23):8831–8838, 1991.
40. P. F. Stadler, W. Fontana, and J. H. Miller. Random catalytic reaction networks. *Physica D*, 63:378–392, 1993.

41. P. F. Stadler and R. Happel. The probability for permanence. *Math. Biosc.*, 113:25–50, 1993.
42. P. F. Stadler and P. Schuster. Mutation in autocatalytic networks - an analysis based on perturbation theory. *J. Math. Biol.*, 30:597–631, 1992.
43. E. Szathmáry and I. Gladkih. Sub-exponential growth and coexistence of non-enzymatically replicating templates. *J. Theor. Biol.*, 138:55–58, 1989.
44. T. Tjivikua, P. Ballester, and J. Rebek Jr. A self-replicating system. *J. Am. Chem. Soc.*, 112:1249–1250, 1990.
45. O. C. Uhlenbeck. A small catalytic oligoribonucleotide. *Nature*, 328: 596–600, 1987.
46. G. von Kiedrowski. A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed. Engl.*, 25:932–935, 1986.
47. 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.
48. G. von Kiedrowski, B. Wlotzka, and J. Helbing. Sequence dependence of template-directed syntheses of hexadeoxynucleotide derivatives with 3'-5' pyrophosphate linkage. *J. Angew. Chem. Int. Edn. engl.*, 28:1235–1237, 1989.
49. P. R. Wills, S. A. Kauffman, B. M. R. Stadler, and P. F. Stadler. Selection dynamics in autocatalytic systems: Templates replicating through binary ligation. *Bull. Math. Biol.*, 60:1073–1098, 1998.
50. B. Wlotzka and J. S. McCaskill. A molecular predator and its prey: Coupled isothermal amplification of nucleic acids. *Chemistry & Biology*, 4:25–33, 1997.
51. S. Yao, I. Ghosh, R. Zutshi, and J. Chmielewski. Selective amplification by auto- and cross-catalysis in a replicating peptide system. *Nature*, 396:447–450, 1998.