Selection Dynamics in Autocatalytic Systems: Templates Replicating Through Binary Ligation

Peter R. Wills[a,b,*], Stuart A. Kauffman[a], Bärbel M. R. Stadler[c], and Peter F. Stadler[a,d]

[a]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico, 87501, USA
[b]Dept. of Physics, University of Auckland, Private Bag 92019, Auckland, New Zealand
[c]Dept. of Statistics Operations Research and Computer Methods, University of Vienna, Austria
[d]Dept. of Theoretical Chemistry, University of Vienna, Austria

Address for correspondence: Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, phone (505) 984 8800 ext. 259, fax: (505) 982 0565, email wills@santafe.edu

Abstract

The theory of autocatalytic binary ligation is reviewed within the context of a consistently applied Michaelis-Menten quasi-steady-state approximation to obtain explicit analytical results describing time-course data from experiments. A detailed protocol for the step-wise elucidation of a minimal set of experimental parameters is outlined. The kinetic equations are then generalized to cases of self- and cross-catalysis among an arbitrary number of different templates and applied to experiments involving just two templates. Depending on the values of various kinetic parameters such systems can display exclusionary Darwinian selection corresponding to an exponential growth law, selective coexistence or coexistence of all species characteristic of a parabolic growth law; the intermediate behaviour arises as a property of the full mechanism analysed here. Our results are applicable to the classical case of self-replicating nucleic acids and their analogues as well as to newly discovered self-replicating peptides.

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 [31] and the general description of such systems has been termed “Minimal Replicator Theory” [32]. 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 [17]. Molecular self-replication of this sort is of central importance to our understanding the origin of life and it was within this context that the selection dynamics of template-directed autocatalysis were first discussed [5, 7]. 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 [33, 23], although an example of complex, mixed, auto- and cross-catalysis among sets of closely related templates has been reported [1]. However the 32 residues of the self-replicating peptide of Lee *et al.* [17] offer much greater scope for the creation of sets of closely related templates and reactants whose possibilities for auto- and cross-catalysis are likely to display many modes of selective behavior, depending on details of the chemical kinetic processes whereby template synthesis takes place.

The nucleic acid and peptide templates that have been studied both display a ”parabolic” autocatalytic growth pattern [31, 17, 20, 29, 34]. It is relatively easy to derive a rate equation displaying this elementary behavior if one assumes that catalysis proceeds through the complementary binding of reactant(s) to free template and that auto-catalysis is limited by the tendency of the template to bind to itself as an inactive "product inhibited" dimer [32]. 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. In this paper we show that the usual Michaelis-Menten description of biopolymeric catalysis gives equations of remarkable simplicity in appropriately chosen variables. Even the expressions for some more complicated observable quantities are amenable to direct analytical solution. Our analysis provides a means for the rapid development of detailed kinetic "signatures" whereby the behavior of systems can be characterized in terms of differently structured patterns within the catalytic network underlying the general dynamics. We find that these systems can display evolutionary behavior which is intermediate between competitive exclusion [5] and universal coexistence [28].

2. Background Theory

We begin by showing that when the Michaelis-Menten approximation [19] is cast in a way slightly different from the formulation used in [32], the solution of the differential equations describing template-directed binary ligation can be written in a simpler closed form. Reformulating the Michaelis-Menten approximation also allows us to generalize the analytical treatment to the description of competitive and cooperative systems and thereby comment on the selection dynamics of systems containing mixtures of templates with different catalytic specificities.

Consider the autocatalytic synthesis of a template C through ligation of two fragments A and B. The simplest possible description would be in terms of the chemical equation

$$A + B + C \to 2C \tag{2.1}$$

and the corresponding rate equation

$$\frac{d}{dt}[C] = k[A][B][C]\,, \tag{2.2}$$

which leads to the expectation of initially exponential growth of C if the initial concentration $[C]^\circ$ of C is much less than that of A or B:

$$[C] = [C]^\circ \exp\left(k[A]^\circ[B]^\circ t\right) \tag{2.3}$$

However this description does not represent the observation that the autocatalytic growth of both nucleotide and peptide templates is "parabolic", meaning that the concentration grows with t^2. The minimal model able to explain this observation consists of the three reactions [32]:

$$A + B + C \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ABC\,, \qquad K_1 = k_1/k_{-1} \tag{2.4}$$

$$C_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} 2C\,, \qquad K_2 = k_2/k_{-2} \tag{2.5}$$

$$ABC \xrightarrow{k_3} C_2\,, \tag{2.6}$$

where we have defined the equilibrium constants K_i for reactions (2.4) and (2.5) in terms of rate constants k_i and k_{-i} for the corresponding forward and backward reactions and the essentially irreversible step (reaction 2.6) is governed by a rate constant k_3.

The differential equations expressing the rates of formation of each of the C-containing species are then

$$\frac{d}{dt}[C] = k_{-1}[ABC] - k_1[A][B][C] + 2k_2[C_2] - 2k_{-2}[C]^2 \tag{2.7}$$

$$\frac{d}{dt}[ABC] = k_1[A][B][C] - (k_{-1} + k_3)[ABC] \tag{2.8}$$

$$\frac{d}{dt}[C_2] = k_3[ABC] + k_{-2}[C]^2 - k_2[C_2] \tag{2.9}$$

Denoting the total molar concentration of the constituent C by

$$c = [C] + [ABC] + 2[C_2] \tag{2.10}$$

we obtain from Eqns. (2.7-2.9):

$$\frac{dc}{dt} = k_3[ABC] \tag{2.11}$$

reflecting the fact that ligation of A and B on the template C is the only process which results in a net production of C.

Provided the Segel-Slemrod condition [22, 4] holds, which in this case reads $[A] + (k_{-1} + k_3)/k_1 \gg [A] > 0$, and similarly for $[B]$, we may make the Michaelis-Menten quasi-steady state assumption ($d[ABC]/dt = 0$) in respect of the intermediate ABC and obtain the relation

$$[ABC] = \frac{k_1}{(k_{-1} + k_3)}[A][B][C]. \tag{2.12}$$

We now apply the Michaelis-Menten assumption to the latter phase of the reaction sequence, setting $d[C_2]/dt = 0$ under the Segel-Slemrod condition $[C] + (k_2 + k_3)/k_{-2} \gg [C] > 0$. Equs.(2.10)-(2.12) then yield the extremely simple result

$$\frac{d}{dt}[C] = \frac{k_1 k_3}{(k_{-1} + k_3)}[A][B][C] \tag{2.13}$$

If $[A]$ and $[B]$ can be regarded as constant, either because the concentrations of A and B are buffered or their depletion due to reaction (2.4) is negligible on some relevant time-scale, then we retrieve equation (2.2) with $k = k_1 k_3/(k_{-1} + k_3)$. Thus, irrespective of the fact that the overall reaction mechanism involves the formation of complexes ABC and C_2, the concentration of free template $[C]$ displays exponential growth and it represents a "natural selection variable". However the exponential form of Equ.(2.13) does not mean that this will be the growth pattern observed by measuring the total constituent concentration defined in Equ.(2.10).

If the predominant form of the template is the free, unbound species ($K_1[A][B] \ll 1 \ll K_2; [C] \gg [ABC] + 2[C_2]$) then we can substitute $c \approx [C]$ in Equ.(2.13) and expect to observe exponential growth when $[A]$ and $[B]$ are effectively constant. On the other hand, if the predominant form of the template is the dimeric species ($K_2 \ll 1; K_1[A][B] \ll 1; [C_2] \gg [C] + [ABC]$) then $c \approx 2[C_2]$ and $[C_2]$ can be expressed reasonably accurately in terms of $[C]$ through the equilibrium condition

$$[C_2] = [C]^2/K_2 \tag{2.14}$$

and we then obtain the differential equation

$$\frac{dc}{dt} = K_1 k_3 \sqrt{K_2/2}[A][B]\sqrt{c} \tag{2.15}$$

With $[A]$ and $[B]$ effectively constant we obtain the solution

$$\sqrt{c} = (\alpha/2)t + \sqrt{c^\circ} \tag{2.16}$$

where $\alpha = K_1 k_3 \sqrt{(K_2/2}[A][B]$. At times for which $\sqrt{c} \gg \sqrt{c^\circ}$, we obtain the parabolic growth law

$$c \approx \frac{\alpha^2}{4}t^2 \tag{2.17}$$

which is identical to von Kiedrowski's Equ.(11) in [32].

As a way of treating cases intermediate between exponential and parabolic growth, von Kiedrowski [32] introduced the semi-empirical equation

$$\frac{dc}{dt} = ([A]^\circ - c)([B]^\circ - c)[k(c + c^\circ)^p + k_b]e^{-k_d t} \tag{2.18}$$

based on considerations of the depletion of substrates A and B, which occurs in the time course of many experiments, as well as the uncatalysed production (with rate constant k_b) and irreversible hydrolysis (with rate constant k_d) of template C due to the reactions

$$A + B \rightarrow C \rightarrow products \tag{2.19}$$

Equ.(2.18) has proved useful for analyzing the kinetics of template replication and understanding the thermodynamic constraints on the process in terms of the reaction order parameter, p, which itself has a complex definition in terms of fundamental rate constants [31]. We will now show that if $[A]$ and $[B]$ can be regarded as effectively constant then it is possible to derive a useful exact solution of the differential equations describing all of the chemical processes involved in template replication. We also show how our analysis can be extended to the understanding of simple experiments involving more than one template, and the general evolutionary dynamics of such systems.

3. Explicit Solution for a Single Template

We begin by defining a Michaelis constant K_m for the autocatalytic enzymic production of template C via reactions (2.4) and (2.6):

$$K_m = \frac{(k_{-1} + k_3)}{k_1} \tag{3.1}$$

Applying the quasi-steady state (Michaelis-Menten) assumption again, this time to the dimeric template ($d[C_2]/dt = 0$) we obtain,

$$[C_2] = \frac{k_3}{K_m k_2}[A][B][C] + \frac{1}{K_2}[C]^2 \tag{3.2}$$

We can then express Equ.(2.10) in the form

$$c = \beta[C] + \frac{2}{K_2}[C]^2 \tag{3.3}$$

where

$$\beta = 1 + \frac{1}{K_m}(1 + \frac{2k_3}{k_2})[A][B] \tag{3.4}$$

Solving for the free concentration of template, $[C]$, in terms of the total concentration, c, gives

$$[C] = \frac{\sqrt{\beta^2 + 8c/K_2} - \beta}{4/K_2} \tag{3.5}$$

which can be substituted into Equ.(2.13) to separate the variables of integration:

$$\frac{4K_m}{k_3 K_2 [A][B](\sqrt{\beta^2 + 8c/K_2} - \beta)} dc = dt \tag{3.6}$$

For $[A]$ and $[B]$ constant, we integrate to obtain the explicit result

$$\frac{K_m}{k_3[A]^\circ[B]^\circ}[Y - Y^\circ + \beta \ln\frac{(Y - \beta)}{(Y^\circ - \beta)}] = t \tag{3.7}$$

where

$$Y = \sqrt{\beta^2 + 8c/K_2} \tag{3.8}$$

Although t is the subject of Equ.(3.7) it can be used for the direct fitting of multiple experimental results of the time-course of template replication reactions (c versus t) at different values of $[A]^\circ$, $[B]^\circ$ and c° in terms of three unknown parameters, K_2 , K_m/k_3 and ϵ where

$$\epsilon = \frac{1}{K_m}(1 + \frac{2k_3}{k_2}) = \frac{\beta - 1}{[A][B]} \tag{3.9}$$

If we wish to take into account the non-catalytic processes described by reaction (2.19), then we need a modified version of Equ.(2.13):

$$\frac{d}{dt}[C] = \frac{dc}{dt} = \frac{k_3}{K_m}[A][B][C] + k_b[A][B] - k_d[C] \tag{3.10}$$

The solution to this equation is

$$\frac{1}{\gamma}[Y - Y^\circ + \delta \ln \frac{(Y - \delta)}{(Y^\circ - \delta)}] = t \tag{3.11}$$

with the coefficients

$$\begin{aligned} \gamma &= \frac{k_3[A][B]}{K_m} - k_d \\ \delta &= \beta - \frac{4k_b[A][B]}{\gamma K_2} \end{aligned} \tag{3.12}$$

Provided k_b and k_d can be determined independently for the relevant experimental conditions, results can be fitted directly to Equ.(3.11) in terms of the same three undetermined parameters relevant to Eqn. (3.7).

If we also wish to take account of the depletion of the substrates A and B, then we must use the relations

$$[A]^\circ - [A] = [B]^\circ - [B] = c - c^\circ \tag{3.13}$$

Even use of the approximation

$$[A][B] \approx [A]^\circ[B]^\circ - ([A]^\circ + [B]^\circ)(c - c^\circ) \tag{3.14}$$

where the product $[A][B]$ appears as a free factor in Equ. (3.7), and ignoring the need to consider any temporal change in β on the justifiable grounds that $\beta \approx 1$, integration yields unwieldy expressions relating c and t. We conclude that reaction time-course data (c *versus* t) from experiments in which substrates are significantly depleted should be fitted to curves generated by numerical integration of Equ. (3.6) or its more general form

$$dc = \{(\gamma K_2/4)(\sqrt{\beta^2 + 8c/K_2} - \beta) + k_b[A][B]\}dt \tag{3.15}$$

to find estimates of the three parameters K_2, K_m/k_3 (related to γ through Equ.(3.12)) and ϵ (related to β through Equ.(3.9)). Direct fitting of experimental data to Equ.(3.15) is computationally more economical than simulation of a reaction model using coupled differential equations as a way of estimating the minimal set of rate constants needed to characterize the dynamics of the system.

Our analysis so far serves to illustrate that when the quasi-steady state condition is applied to the differential equations describing the complete dynamics of replicating template system, the concentration of *free* template $[C]$ follows a simple exponential growth law, Equ. (2.13), irrespective of the behavior observed for the *total* constituent concentration c, provided $[A]$ and $[B]$ are effectively constant and uncatalysed production and degradation can be neglected. An initial exponential increase in c may not always be experimentally observable even though $[C]$ follows such a growth law. We will now show how a similar simple result can be obtained in the much more complex case of many interdependent template-directed replication reactions. This allows for an analytical description of these general systems in which the growth of individual species may arise from processes whose dynamics are exponential, parabolic, intermediate between these two extremes, or a complicated mixture of more than one or all of these possibilities.

4. Multiple Templates

A change of nomenclature is needed to simplify the definition of rate constants when there are multiple templates. Consider the reaction scheme

$$A_i + B_j + C_{kl} \underset{\bar{a}_{ijkl}}{\overset{a_{ijkl}}{\rightleftharpoons}} A_iB_jC_{kl} \overset{b_{ijkl}}{\longrightarrow} C_{ij}C_{kl} \underset{\bar{d}_{ijkl}}{\overset{d_{ijkl}}{\rightleftharpoons}} C_{ij} + C_{kl} \tag{4.1}$$

together with a slow background reaction $A_i + B_j \overset{g_{ij}}{\longrightarrow} C_{ij}$. We have designated the rate constants for the forward reactions by a, b, d, missing out c which stands for concentration. The corresponding reverse reaction rate constants are designated $\bar{a}$, etc. For convenience the template molecules are given double indices such as ij indicating that they are formed by ligating the smaller substrates A_i and B_j.

Assuming that the concentrations of the intermediates are stationary we find

$$\begin{aligned} 0 &= \frac{d}{dt}[A_iB_jC_{kl}] = a_{ijkl}[A_i][B_j][C_{kl}] - (b_{ijkl} + \bar{a}_{ijkl})[A_iB_jC_{kl}] \\ 0 &= \frac{d}{dt}[C_{ij}C_{kl}] = b_{ijkl}[A_iB_jC_{kl}] + b_{klij}[A_kB_lC_{ij}](ij \neq kl) + \bar{d}_{ijkl}[C_{ij}][C_{kl}] - d_{ijkl}[C_{ij}C_{kl}] \end{aligned} \tag{4.2}$$

where the symbol $(ij \neq kl)$ equals 1 if $ij \neq kl$ and 0 if $ij = kl$. Hence

$$\begin{aligned} [A_iB_jC_{kl}] &= \frac{a_{ijkl}}{b_{ijkl} + \bar{a}_{ijkl}}[A_i][B_j][C_{kl}] \\ [C_{ij}C_{kl}] &= \frac{b_{ijkl}a_{ijkl}}{d_{ijkl}(b_{ijkl} + \bar{a}_{ijkl})}[A_i][B_j][C_{kl}] + \frac{b_{klij}a_{klij}}{d_{ijkl}(b_{klij} + \bar{a}_{ijkl})}[A_k][B_l][C_{ij}](ij \neq kl) + \frac{\bar{d}_{ijkl}}{d_{ijkl}}[C_{ij}][C_{kl}] \end{aligned} \tag{4.3}$$

We consider $C_{ij}C_{kl}$ and $C_{kl}C_{ij}$ as equivalent notations for the same chemical species. Hence the "diagonal terms", $ij = kl$, occur only once in Eqns.(4.2) and (4.3).

The dynamics for the free reaction product, C_{kl}, is:

$$\begin{aligned} \frac{d}{dt}[C_{kl}] = &-\sum_{ij} a_{ijkl}[A_i][B_j][C_{kl}] + \sum_{ij} \bar{a}_{ijkl}[A_iB_jC_{kl}] \\ &+ \sum_{ij} d_{ijkl}[C_{ij}C_{kl}] + d_{klkl}[C_{kl}C_{kl}] - \sum_{ij} \bar{d}_{ijkl}[C_{ij}][C_{kl}] - \bar{d}_{klkl}[C_{kl}][C_{kl}] \end{aligned} \tag{4.4}$$

Substituting the stationary concentrations of the intermediates we obtain

$$\begin{aligned} \frac{d}{dt}[C_{kl}] = &-\sum_{ij} a_{ijkl}[A_i][B_j][C_{kl}] + \sum_{ij} \frac{\bar{a}_{ijkl}a_{ijkl}}{b_{ijkl} + \bar{a}_{ijkl}}[A_i][B_j][C_{kl}] + \sum_{ij} \frac{b_{ijkl}a_{ijkl}}{b_{ijkl} + \bar{a}_{ijkl}}[A_i][B_j][C_{kl}] \\ &+ \sum_{ij \neq kl} \frac{b_{klij}a_{klij}}{b_{klij} + \bar{a}_{klij}}[A_k][B_l][C_{ij}] + \sum_{ij} \bar{d}_{ijkl}[C_{ij}][C_{kl}] - \sum_{ij} \bar{d}_{ijkl}[C_{ij}][C_{kl}] \\ &+ \frac{b_{klkl}a_{klkl}}{b_{klkl} + \bar{a}_{klkl}}[A_k][B_l][C_{kl}] + \bar{d}_{klkl}[C_{kl}]^2 - \bar{d}_{klkl}[C_{kl}]^2 + g_{kl}[A_k][B_l] \\ = \; & [C_{kl}] \sum_{ij} \left\{ -a_{ijkl} + \frac{\bar{a}_{ijkl}a_{ijkl}}{b_{ijkl} + \bar{a}_{ijkl}} + \frac{b_{ijkl}a_{ijkl}}{b_{ijkl} + \bar{a}_{ijkl}} \right\} [A_i][B_j] \\ &+ \sum_{ij} \frac{b_{klij}a_{klij}}{b_{klij} + \bar{a}_{klij}}[A_k][B_l][C_{ij}] + g_{kl}[A_k][B_l] \end{aligned} \tag{4.5}$$

The coefficients in the curly bracket cancel. (Note that by definition $d_{ijkl} = d_{klij}$.) It will be convenient to introduce the inverse Michealis-Menten-type constants

$$m_{klij} = \frac{b_{klij}a_{klij}}{b_{klij} + \bar{a}_{klij}} \tag{4.6}$$

and the abbreviations $\tilde{m}_{klij} = [A_k][B_l]m_{klij}$ and $\tilde{g}_{kl} = [A_k][B_l]g_{kl}$. Specific degradation of templates can be introduced easily. We simply add the degradation term $-\Delta_{kl}[C_{kl}]$ to equ.(4.5). With this notation Equ.(4.6) becomes

$$\frac{d}{dt}[C_{kl}] = \sum_{ij} \tilde{m}_{klij}[C_{ij}] - \Delta_{kl}[C_{kl}] + \tilde{g}_{kl} \tag{4.7}$$

A detailed analysis of the behavior of the multi-template system (4.7) is feasible only if we assume that the building-block concentrations $[A_k]$ and $[B_l]$ are buffered. In this case Equ.(4.7) is a linear ordinary differential equation for which explicit solutions can be obtained, see e.g., [10].

The total concentration of a template species C_{ij} is

$$\begin{aligned} c_{kl} &= [C_{kl}] + \sum_{ij}[A_iB_jC_{kl}] + \sum_{ij}[C_{ij}C_{kl}] + [C_{kl}C_{kl}] \\ &= [C_{kl}] + [C_{kl}]\sum_{ij}\frac{a_{ijkl}}{b_{ijkl}+\bar{a}_{ijkl}}[A_i][B_j] + [C_{kl}]\sum_{ij}\frac{b_{ijkl}a_{ijkl}}{d_{ijkl}(b_{ijkl}+\bar{a}_{ijkl})}[A_i][B_j] \\ &\quad + \sum_{ij\neq kl}\frac{\bar{d}_{ijkl}}{d_{ijkl}}[C_{ij}][C_{kl}] + 2\frac{\bar{d}_{klkl}}{d_{klkl}}[C_{kl}]^2 + \sum_{ij}\frac{b_{klij}a_{klij}}{d_{ijkl}(b_{klij}+\bar{a}_{ijkl})}[A_k][B_l][C_{ij}] \\ &= [C_{kl}]\left\{1 + \sum_{ij}\frac{\tilde{m}_{ijlk}}{b_{ijkl}} + \sum_{ij}\frac{\tilde{m}_{ijlk}}{d_{ijkl}}\right\} + \sum_{ij\neq kl}\frac{\bar{d}_{ijkl}}{d_{ijkl}}[C_{ij}][C_{kl}] + 2\frac{\bar{d}_{klkl}}{d_{klkl}}[C_{kl}]^2 + \sum_{ij}\frac{\tilde{m}_{klij}}{d_{ijkl}}[C_{ij}] \end{aligned} \tag{4.8}$$

In general it is not possible to obtain explicit expressions for the concentrations of the free templates, $[C_{ij}]$, in terms of the total template concentrations, c_{ij}. Equ.(4.7) is realistic only for relatively short time scales, since the total template concentration

$$c = \sum_{ij} c_{ij} \tag{4.9}$$

increases exponentially. Hence, a more realistic model must include some control of the total concentration. To this end we introduce the relative concentrations $x_{ij} = c_{ij}/c$.

From the first line of Equ.(4.8) and the stationarity assumption for all intermediates we find $dc_{kl}/dt = d[C_{kl}]/dt$. A short computation then yields

$$\begin{aligned} \dot{x}_{kl} &= \dot{c}_{kl}/\sum_{rs} c_{rs} - x_{kl}\sum_{ij}\dot{c}_{ij}/\sum_{rs} c_{rs} = \frac{1}{c}\left(\frac{d[C_{kl}]}{dt} - x_{kl}\sum_{ij}\frac{d[C_{ij}]}{dt}\right) \\ &= \frac{1}{c}\left(\sum_{ij}\tilde{m}_{klij}[C_{ij}] - \Delta_{kl}[C_{kl}] - x_{kl}\left(\sum_{ijrs}\tilde{m}_{rsij}[C_{ij}] - \sum_{ij}\Delta_{ij}[C_{ij}]\right)\right) + \frac{1}{c}\left(\tilde{g}_{kl} - x_{kl}\sum_{rs}\tilde{g}_{rs}\right) \end{aligned} \tag{4.10}$$

The degradation term in Equ.(4.7) affects the diagonal terms only. A simple modification $\tilde{m}_{klkl} \to \tilde{m}_{klkl} - \Delta_{kl}$ in Equ.(4.10) will take care of its effect. Just as in Eigen's quasi-species model it affects the net fitness but does not change the qualitative behavior of the system. Since we shall not attempt quantitative predictions in this (and the following) section we will not explicitly display the degradation terms in the following.

As long as the total concentration satisfies $0 < D < c(t) < M < \infty$ with some positive constants D and M we may rescale the time axis and simply drop the factor of $1/c$ without changing the phase portraits [12]. Note that Equ.(4.10) is valid also in the presence of an *unspecific dilution flux* depleting each molecular species proportional to its concentration, since such a flux has no influence on the *relative* abundances x_{kl}. The will of course determine the time-course of the total concentration c of templates.

The non-catalyzed production $\tilde{g}_{kl} - x_{kl}\overline{\tilde{g}} > 0$ if x_{kl} is sufficiently small and if $\tilde{g}_{kl} > 0$. Thus, if $[C_{ij}]$ is sufficiently small we have $\dot{x}_{ij} > 0$, i.e., the dynamical system Equ.(4.10) always leads to coexistence. It is not surprising, of course, that all templates coexist if they can be formed spontaneously.

5. Analysis of Some Simple Experiments

The results of Section 4 can be applied to the analysis of experiments which have been conducted using more than one substrate of type A or B, or more than one template. We consider first the situation in which the elementary reaction $A_1 + B_1 \to C_{11}$ occurs in the presence of an additional template, C_{12} say. In this situation we can simplify the nomenclature by dropping the subscripts on A and B and using $C_{11} \to C_1$, $C_{12} \to C_2$, and correspondingly, $a_{1111} \to a_1$, $b_{1112} \to b_2$, $d_{1111} \to d_{11}$, etc. Solution of Equ.(4.8) for template 1 followed by differentiation and substitution gives an analogue of Equ.(3.14):

$$dc_1 = \{(\gamma d_{11}/4\bar{d}_{11})(\sqrt{\beta^2 + 8\bar{d}_{11}(c_1 - \zeta)/d_{11}} - \beta) + \tilde{m}_2[C_2] + \tilde{g}_1\}dt \tag{5.1}$$

where $\gamma = \tilde{m}_1 - \Delta_1$ as defined in Equ.(3.11) and we now have

$$\beta = 1 + \epsilon[A][B] + (\bar{d}_{12}/d_{12})[C_2] \tag{5.2}$$

with $\epsilon = (m_1/b_1)(1 + 2b_1/d_{11})$ as defined in Equ.(3.15) and

$$\zeta = \frac{\tilde{m}_1}{d_{12}}[C_2] \tag{5.3}$$

First it should be noted that the parameters $d_{11}/\bar{d}_{11}$, γ and ϵ can all be determined from independent, separate experiments using the single replicating templates C_1 and C_2. Besides $\tilde{m}_2$ that leaves β and ζ, both linear functions of $[C_2]$, to be determined from c_1 *versus* t time-course data at different values of c_2, the concentration of the second template. The question to be answered is whether the free concentration $[C_2]$ can be expressed as a function of the total concentration c_2 under any reasonable approximation. It would be ideal if an approximation could be found under which $[C_2]$ could be considered constant during a time-course measurement so that Equ.(5.3) can be integrated numerically.

For the added template C_2, solution of Equ.(4.8) yields

$$[C_2] = \frac{\sqrt{\beta'^2 + 8(\bar{d}_{22}/d_{22})c_2} - \beta'}{4(\bar{d}_{22}/d_{22})c_2} \tag{5.4}$$

$$\text{where} \qquad \beta' = 1 + \epsilon'[A][B] + (\bar{d}_{12}/d_{12})[C_1] \tag{5.5}$$

$$\text{and} \qquad \epsilon' = \frac{m_2}{b_2}(1 + \frac{b_2}{d_{12}})\,. \tag{5.6}$$

There is a hierarchy of approximations to be considered. If the initial concentration of template 1 is zero, then for sufficiently short times the last term in Equ.(5.2) can be neglected and $[A]$ and $[B]$ can be considered constant at their initial values. This condition, which is valid as long as $(\bar{d}_{12}/d_{12})[C_1] \ll 1$, is sufficient for the effective decoupling of the differential equations for the time evolution of C_1 and C_2 and for β and ζ to be determined as functions of c_2. The way in which these functions can be represented depends on what further approximations can be made. For $b_2 \ll d_{12}$, we can use $\beta' \approx 1 + (\tilde{m}_1/b_2)$, and eventually $\beta' \approx 1$ if the Michaelis constant for catalysis using template 2 is also sufficiently large. If the additional condition $d_{22}/\bar{d}_{22} \ll 1$ is satisfied, we can use the approximation $[C_2] \approx \sqrt{d_{22}/(2\bar{d}_{22})}\sqrt{c_2}$.

Our final consideration is the time variation of c_2. Equ.(4.7) gives

$$\frac{dc_2}{dt} = -\Delta_2[C_2] \tag{5.7}$$

which can be integrated directly under the most stringent approximation to yield

$$\sqrt{c_2} \approx \sqrt{c_2{}^\circ} - \Delta_2\sqrt{2d_{22}/\bar{d}_{22}}\; t \tag{5.8}$$

Thus, for sufficiently short times, $t \ll \sqrt{c_2{}^\circ \bar{d}_{22}/(2d_{22})}/\Delta_2$, c_2 and $[C_2]$ can be considered constant. Then the simplest approach to the analysis of experiments of this sort is to fit c_1 *versus* t data to Equ.(5.1) and thereby determine how β and ζ vary with c_2. This variation is likely to be only poorly determined and one should expect at best to detect the dependence of β and ζ on $\sqrt{c_2{}^\circ}$ to obtain estimates of $(\tilde{m}_1/d_{12})$ and $(\bar{d}_{12}/d_{12})$.

The second set of experiments to be considered are those in which both templates can be synthesized from different substrates via reactions like $A+B_1 \to C_1$ and $A+B_2 \to C_2$ and both C_1 and C_2 can act as catalysts for either reaction. This situation is significantly more complex than that in which just a single template is synthesized, but it is still tractable. We note first that the indices in the equations of Section 4 can once again be simplified through the reduction $ijkl \to jl$.

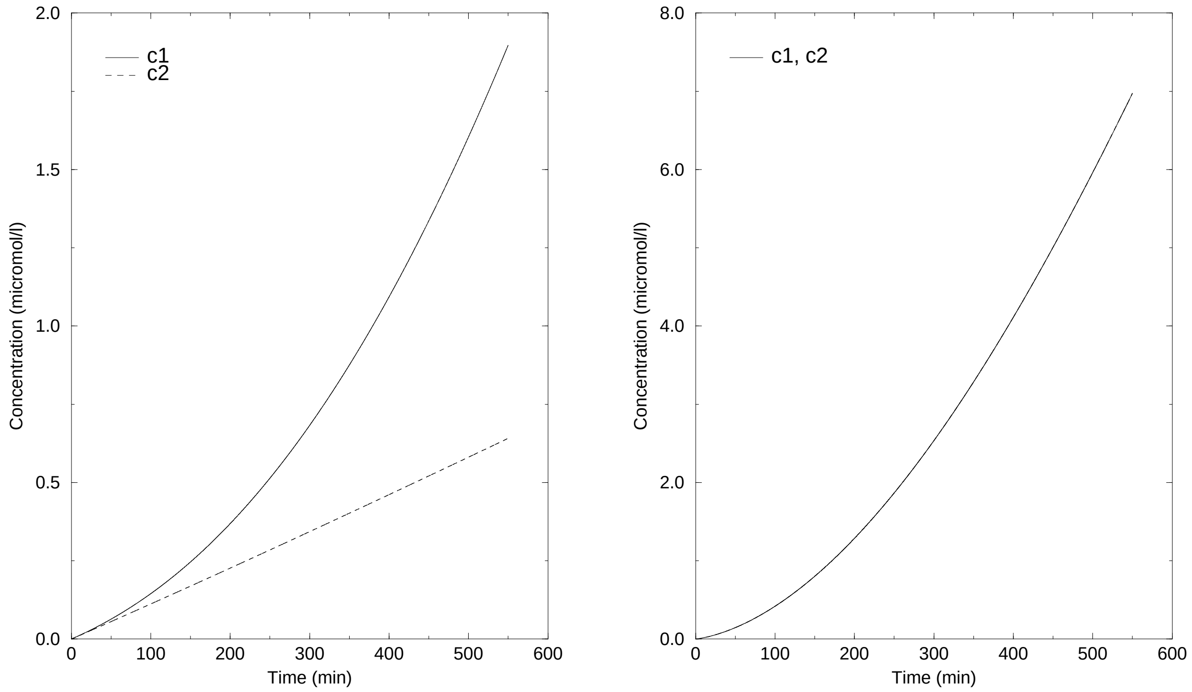

Figure 1. Examples of time course data for the production of two templates in accordance with equ.(5.1) under conditions of Darwinian competition and mutual coexistence. Parameters have been chosen to approximate experiments like those of [17]. In each case $\Delta_i = 0$, $g_i = 1.14 \cdot 10^{-3}\mathrm{M}^{-1}\mathrm{s}^{-1}$, $a_{ij} = 10^7\mathrm{M}^{-2}\mathrm{s}^{-1}$, $\bar{a}_{11} = 3 \cdot 10^{-3}\mathrm{s}^{-1}$, $\bar{a}_{12} = \bar{a}_{21} = 5 \cdot 10^{-4}\mathrm{s}^{-1}$, $\bar{a}_{22} = 10^{-4}\mathrm{s}^{-1}$, $\bar{d}_{ij} = 6 \cdot 10^9\mathrm{M}^{-1}\mathrm{s}^{-1}$, $[A]^\circ = 1.8 \cdot 10^{-4}\mathrm{M}$, and $[B_i]^\circ = 9 \cdot 10^{-5}\mathrm{M}$, $c_i^\circ = 0$. **(a)** Competition. $b_{11} = 10^4\mathrm{s}^{-1}$, $b_{12} = 30\mathrm{s}^{-1}$, $b_{21} = 5\mathrm{s}^{-1}$, $b_{22} = 300\mathrm{s}^{-1}$, $d_{11} = 3.5 \cdot 10^{-3}\mathrm{s}^{-1}$, $d_{12} = d_{21} = 10^{-4}\mathrm{s}^{-1}$, $d_{22} = 10^{-5}\mathrm{s}^{-1}$. **(b)** Coexistence. $b_{11} = 300\mathrm{s}^{-1}$, $b_{12} = 2000\mathrm{s}^{-1}$, $b_{21} = 8000\mathrm{s}^{-1}$, $b_{22} = 100\mathrm{s}^{-1}$, $d_{ij} = 1 \cdot 10^{-2}\mathrm{s}^{-1}$.

For template 1 the time evolution is given by an equation of the same form as Equ.(5.1) with

$$\beta_1 = 1 + \epsilon_1[A][B_1] + \epsilon_1'[A][B_2] + (\bar{d}_{12}/d_{12})[C_2] \tag{5.9}$$

and, as previously defined, $\gamma_1 = \tilde{m}_{11} - \Delta_1$, $\zeta_1 = (\tilde{m}_{11}/d_{12})[C_2]$, $\epsilon_1 = (m_{11}/b_{11})(1 + 2b_{11}/d_{11})$ and $\epsilon_1' = (m_{21}/b_{21})(1 + b_{21}/d_{21})$. The equation for the time evolution of template 2 is identical with complete substitution of all indices $2 \to 1$ and $1 \to 2$. Note first of all that the parameters γ_1, γ_2, $(\tilde{m}_{11}/d_{12})$, $(\tilde{m}_{22}/d_{12})$, ϵ_1, ϵ_2, $\tilde{m}_{12}$ and $(\bar{d}_{12}/d_{12})$ can all be determined in preliminary experiments of the sort described above. That leaves ϵ_1' and ϵ_2', which cannot be determined effectively in preliminary experiments, as free parameters. The main problem that arises is that the coupling between the equations for the time evolution of c_1 and c_2 is expressed irreducibly in terms of $[C_1]$ and $[C_2]$. Furthermore, $[C_1]$ and $[C_2]$ cannot be determined from c_1

and c_2 without knowledge of ϵ'_1 and ϵ'_2. It would therefore seem that the numerical integration of coupled equations of the form of Equ.(5.1) will have to be performed iteratively in order to obtain estimates of ϵ'_1 and ϵ'_2 from $(c_1,\ c_2)$ *versus* t time-course data.

Some typical examples of the form of results that could be obtained in realistic experiments are shown in Figure 1. The particular parameters have been chosen to illustrate the diversity of dynamic behavior possible in such systems. This implies the need to follow a stepwise procedure, such as that outlined in Eqns.(5.1-5.9), in elucidating the mode of mutual template growth.

6. Coexistence and Competition of Multiple Templates

In this section we shall discuss the long time behavior of Equ.(4.10) in the absence of the spontaneous production terms. We shall consider two extremal cases:
(1) The total concentration c is small. Then we may neglect the quadratic terms in Equ.(4.8).
(2) The total concentration c is large, then the quadratic terms will dominate in Equ.(4.8).

In the first case $c_{kl} = \sum_{ij} \Theta_{kl,ij}[C_{ij}]$, where $\mathbf{\Theta}$ is an invertible matrix. Thus

$$\frac{[C_{kl}]}{c} = \sum_{ij} \Theta^{-1}_{kl,ij} x_{ij} \tag{6.1}$$

where the coefficients $\Theta^{-1}_{kl,ij}$ are the entries of the inverse of $\mathbf{\Theta}$. Note that the degradation constants Δ_{kl} do not enter Equ.(4.8); hence $\mathbf{\Theta}$ is also independent of the degradation rates. We obtain

$$\dot{x}_{kl} = \sum_{ij} \tilde{m}_{klij} \sum_{rs} \Theta^{-1}_{ij,rs} x_{rs} - x_{kl} \sum_{pq} \sum_{ij} \tilde{m}_{pqij} \sum_{rs} \Theta^{-1}_{ij,rs} x_{rs} \tag{6.2}$$

It pays to introduce some more notation here: Denote the matrix of Michaelis-Menten-type elements [see Equ.(4.6)] by $\tilde{\mathbf{M}}$ and define the *fitness* of template C_{ij}

$$\alpha_{ij} = \sum_{rs} [\tilde{\mathbf{M}}\mathbf{\Theta}^{-1}]_{rsij} \tag{6.3}$$

and set $Q_{klij} = [\tilde{\mathbf{M}}\mathbf{\Theta}^{-1}]_{rsij}/\alpha_{ij}$. Then Equ.(6.2) becomes

$$\dot{x}_{kl} = \sum_{ij} Q_{klij}\alpha_{ij} - x_{kl} \sum_{ij} \alpha_{ij} x_{ij} \tag{6.4}$$

This is Eigen's quasi-species equation [5]. The off-diagonal elements of the "mutation rate matrix" $\mathbf{Q}$ vanish if (and only if) the matrix $\tilde{\mathbf{M}}$ is diagonal, i.e., if each template instructs only its own production. It makes perfect sense, therefore, to consider the off-diagonal entries of $\mathbf{Q}$ as mutation rates. Since $\dot{x}_{kl} \geq 0$ whenever $x_{kl} = 0$ (the simplex of relative concentrations must be forward invariant), we conclude that $Q_{klij} \geq 0$ if $ij \neq kl$. It is not hard to check that the diagonal is non-negative as well, at least in the absence of degradation terms.

The quasi-species equation yields a unique, globally stable equilibrium provided the non-negative matrix $\mathbf{Q}$ is irreducible. (This is an immediate consequence of the Perron-Frobenius theorem, see [6] for the details).

Let us now consider the other extremal case. Equ.(4.8) can be approximated by neglecting the linear terms and we are left with

$$c_{kl} = \sum_{ij\neq kl} \frac{\bar{d}_{ijkl}}{d_{ijkl}}[C_{ij}][C_{kl}] + 2\frac{\bar{d}_{klkl}}{d_{klkl}}[C_{kl}]^2 \tag{6.5}$$

Note that the template concentration is dominated by the duplex formation in this limit. This quadratic equation still does not allow for an explicit solution.

The situation becomes tractable, however, if we assume that the stability of hetero-duplexes is much smaller than the stability of the homo-duplex. Equ.(6.5) then simplifies to

$$c_{kl} = 2\frac{\bar{d}_{klkl}}{d_{klkl}}[C_{kl}]^2 \tag{6.6}$$

and hence $[C_{kl}] = \eta_{kl}\sqrt{c_{kl}}$, where $\eta_{kl} = \sqrt{d_{klkl}/(2\bar{d}_{klkl})} > 0$. Substituting this into Equ.(4.10) and dropping a factor $1/\sqrt{c}$ (by means of rescaling of the time axis) yields

$$\dot{x}_{kl} = \sum_{ij} \tilde{m}_{klij}\eta_{ij}\sqrt{x_{ij}} - x_{kl}\sum_{rs,ij} \tilde{m}_{rsij}\eta_{ij}\sqrt{x_{ij}} \tag{6.7}$$

This equation can be regarded as a mutation-variant of Szathmáry's model; see the following section. If $\tilde{\mathbf{M}}$ is irreducible, then the boundary of the simplex is a (strict) repellor, and the model leads to cooperation. This has also been observed in other autocatalytic networks with mutation; see e.g., [11, 27, 26].

Very similar effects are observed in the presence of a slow background reaction, $\tilde{g}_{kl} > 0$ in Equ.(4.7), as demonstrated in Figure 1. Figure 1(a) shows that a species with a simple replication advantage rapidly outgrows another whose autocatalytic production is negligible (nevertheless, the inferior species is not entirely removed from the system because of the small linear growth term). Figure 1(b) shows that when reproduction is quasi-complementary, that is, when the main catalyst for each template is the other template, the rates of growth match each other almost exactly. In this case the replicating entity is essentially the template pair [25].

7. Correct Instruction

Correct instruction means $m_{klij} = \tilde{m}_{klij} = Q_{klij} = 0$ for $kl \neq ij$. Equ.(4.8) simplifies to

$$c_{kl} = [C_{kl}]\left\{1 + \frac{\tilde{m}_{klkl}}{b_{klkl}} + \frac{\tilde{m}_{klkl}}{d_{klkl}}\right\} + \sum_{ij\neq kl}\frac{\bar{d}_{ijkl}}{d_{ijkl}}[C_{ij}][C_{kl}] + 2\frac{\bar{d}_{klkl}}{d_{klkl}}[C_{kl}]^2 \tag{7.1}$$

and Equ.(4.10) reduces to

$$\dot{x}_{kl} = \tilde{m}_{klkl}\frac{[C_{kl}]}{c} - x_{kl}\sum_{ij}\tilde{m}_{ijij}\frac{[C_{ij}]}{c} \tag{7.2}$$

Again, there is no analytical solution to Equ.(7.1). In the case of a small total concentration c we find

$$c_{kl} = [C_{kl}]\left\{1 + \frac{\tilde{m}_{klkl}}{b_{klkl}} + \frac{\tilde{m}_{klkl}}{d_{klkl}}\right\} = \Theta_{kl}[C_{kl}] \tag{7.3}$$

and we arrive at a the first order replicator equation

$$\dot{x}_{kl} = x_{kl}\left(\tilde{m}_{klkl}\Theta_{kl}^{-1} - \sum_{ij}\tilde{m}_{ijij}\Theta_{ij}^{-1}x_{ij}\right) \tag{7.4}$$

This model leads to competitive exclusion, that is, in the long time limit there is only a single surviving species, namely the one with the largest coefficient $\tilde{m}_{klkl}\Theta_{kl}^{-1}$. Equ.(7.4) is the zero-mutation limit of Eigen's quasispecies equation.

The appearance of a square root in Eqns.(2.15) and (3.5) suggests the ansatz $[C_{kl}] \sim c_{kl}^{p_{kl}}$ with an effective reaction order p_{kl} that might depend on the parameters and that might be different from 1/2 or 1. A short computation yields

$$\frac{[C_{kl}]}{c_{kl}} \frac{\partial c_{kl}}{\partial [C_{kl}]} = \frac{1}{p_{ij}} \tag{7.5}$$

under the assumption that the variation of p_{ij} with $[C_{ij}]$ is negligible. Writing $\Psi_{ij}(\,.\,)$ as an abbreviation for the r.h.s. of Equ.(7.1) except for the quadratic term, and setting $K_{kl} = \bar{d}_{klkl}/d_{klkl}$ for the stability constant of the homo-duplex $C_{kl}C_{kl}$ we find

$$p_{kl} = \frac{\Psi_{kl}(\,.\,) + 2K_{kl}[C_{kl}]}{\Psi_{kl}(\,.\,) + 4K_{kl}[C_{kl}]}\,. \tag{7.6}$$

If the homo-complexes $[C_{kl}C_{kl}]$ are very stable, i.e., if $K_{kl}[C_{kl}] \gg \Psi_{kl}(\,.\,)$, then we have $p_{kl} = 1/2$. In this case we find $[C_{kl}] \approx \sqrt{K_{kl}c_{kl}/2} = \eta_{kl}\sqrt{c_{kl}}$, and we recover Szathmáry's selection equation

$$\dot{x}_{kl} = \eta_{kl}\tilde{m}_{klkl}x_{kl}^{p} - x_{kl}\sum_{ij}\tilde{m}_{ijij}\eta_{ij}x_{ij}^{p} \tag{7.7}$$

with $p = 1/2$. More directly, Equ.(6.7) reduces to (7.7) is the limit of correct copying. Szathmáry [28] showed that Equ.(7.7) exhibits cooperative behavior. In Appendix A we show that it has a globally stable fixed point in the interior of the state space. In technical terms, Equ.(7.7) is *permanent* [21]. A dynamical system is permanent if, after sufficiently long time, the concentrations of all constituents are bounded from both above and below by two constants that are independent of the concentrations of the constituents and time. For a recent review on permanence see [13].

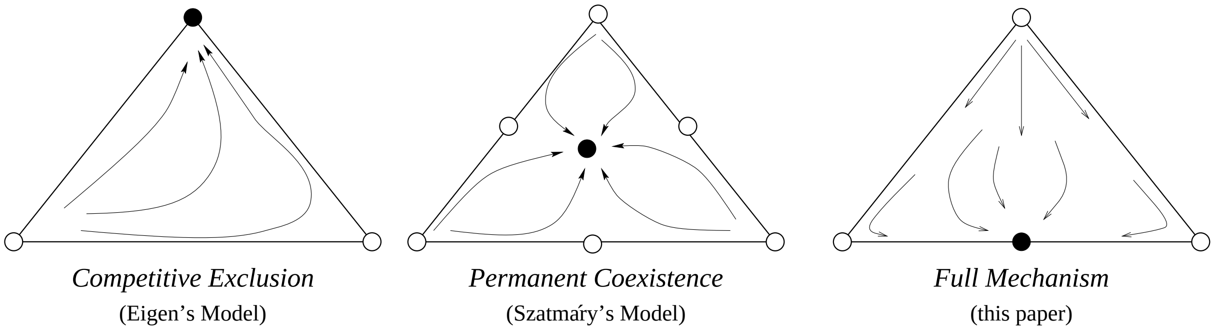

Figure 2. The dynamics of (4.1) is a composite of Eigen's selection model and Szathmáry's model if the homo-complexes are sufficiently stable.

A close inspection of Equ.(7.2) shows, however, that the square-root argument breaks down for $[C_{kl}]$ if c_{kl}, and hence also $[C_{kl}]$ and x_{kl}, is very small. In this case we find $c_{kl} \sim \Psi_{kl}(\,.\,)[C_{kl}]$, i.e., a linear dependence of free and total template concentration.

If the homo-complexes $[C_{kl}C_{kl}]$ are unstable, $K_{kl} = 0$, we find $p_{kl} = 1$ and $[C_{kl}] = c_{kl}/\Psi_{kl}(\,.\,)$ on the entire state space. In this approximation we obtain the generalized replicator equation

$$\dot{x}_{kl} = x_{kl}\left(\varphi_{kl}(\,.\,) - \sum_{ij}x_{ij}\varphi_{ij}(\,.\,)\right) \tag{7.8}$$

where $\varphi_{kl}(\,.\,) = \tilde{m}_{klkl}/\Psi_{kl}(\,.\,)$. A related type of autocatalytic reaction network that reduces to a replicator equation with a non-linear interaction function in some limiting cases is discussed in [9]; see also [24].

A detailed analysis of Equ.(7.8) is not feasible. It is easy, however, to determine the stability of all the 1-species equilibria. The fixed point $x_{kl} = 1$ is stable if all its transversal eigenvectors

$$\Lambda_{pq}(x_{kl} = 1) = \varphi_{pq}(x_{kl} = 1) - \varphi_{kl}(x_{kl} = 1) \tag{7.9}$$

are negative [12]. Note that $x_{kl} = 1$ is equivalent to $[C_{ij}] = 0$ for all $ij \neq kl$. Thus $\Psi_{pq} = f_{pq} + K_{pqkl}[C_{kl}]$. Since we have obtained the replicator equation be assuming that K_{pqpq} is negligible we obtain

$$\Lambda_{pq}(x_{kl} = 1) = \tilde{m}_{pqpq} f_{pq} - q_{klkl} f_{kl} + \tilde{m}_{pqpq} K_{pqkl}[C_{kl}] \qquad kl \neq pq \tag{7.10}$$

Thus there is a single stable 1-species equilibrium determined by $\max(\tilde{m}_{pq} f_{pq})$ if the concentrations of the free templates are small or if the hetero-complexes are also unstable. In the latter case Ψ_{ij} becomes constant and we recover the first-order replicator. Hence we expect *survival of the fittest*.

On the other hand, if $[C_{kl}]$ becomes large (and if the stability constants K_{ijkl} of hetero-complexes are non-negligible) then the $[C_{kl}]$-dependent term may dominate the difference $f_{pq} - f_{kl}$, and we find coexistence of some species or even permanence. Note that f_{pq} is proportional to the concentrations of the unconsumed substrates $[A_p]$ and $[B_q]$. Hence cooperation is more likely to occur if the substrate is kept at a low level while a large total concentration of templates is present.

It is interesting to note that in this case coexistence is not the consequence of a particular network of catalyzed production pathways but originates from the reversible formation of $[C_{ij}C_{kl}]$ complexes. Note, also, that in the above discussion we have nowhere made an explicit assumption on the time-behavior of the total template concentration c, except of course that c is bounded away from 0.

8. Discussion

The demonstration of the first case of a peptide [17] acting autocatalytically to ligate two fragments into a second copy of itself has drawn new attention to a class of reactions which have been studied for some time using nucleic acids [14, 15, 32] and analogues [8]. We are now witnessing important experimental developments in this area. The potential complexity of the reaction systems which can be investigated [18] calls for detailed parallel theoretical work. All of the experimental systems studied so far have been amenable to similar quantitative interpretation on an empirical basis. The main contribution of our current paper is to develop rigorous analytical results which can be used to interpret experimental data in terms of real mechanistic parameters.

(1) By first drawing attention to the long recognized distinction between exponential and parabolic growth in template-directed self-replicating systems we have been able to give a general description of sub-exponential growth in such systems, the detailed form of which depends on the relative affinities of the templates for one another.

(2) We have shown that the complete dynamics of such systems can be represented analytically without any approximation other than the Michaelis-Menten quasi-steady-state assumption. Results from such systems can be analyzed by using simplified numerical integration, without considering the concentration of any of the intermediate species as explicit variables. This is true even when more than one template species is present, and we give explicit examples, both analytical and realistically numerical, for the case of two interacting template species. We have outlined a stepwise procedure for discovering a minimal set of parameters describing the dynamic interaction between two replicating templates. Whether or not the particular prescription of Eqns.(5.1-5.9) is followed depends on what experimental parameters are most easily accessible.

(3) The description of such systems is extended to arbitrary complexity in terms of the number of co-reacting templates and their chemical composition. Whether the binary composition of the template is based on the ligation of peptide or nucleic acid fragments, the mathematical formalism is identical. The distinction between the relatively strict complementarity in the replication of nucleic acids and the much greater diversity of interactions between the peptide templates (which we have called "quasi-complementary") is reflected only in the rate constants assigned to various terms in the differential equations.

(4) In the presence of uncatalyzed template formation (background reaction (2.19)) these systems always lead to permanent coexistence when embedded in a flow reactor. Even in the absence of this background we are bound to find stable coexistence of at least some species under most circumstances: Incorrect instruction

(which corresponds to mutation) forces coexistence, and if the dimers are stable enough the overall dynamics follow a growth law close to parabolic which implies stable coexistence as well [28]. Darwinian selection (survival of the fittest) is observed only when (i) template instruction is sufficiently accurate and (ii) the dimers are sufficiently unstable, or if the total concentration of templates is very small.

(5) The fact that one observes strictly parallel growth in the case of two complementary nucleic acid strands is well recognized [6, 3, 2]. It is now evident that an equivalent situation arises for the much more loosely defined quasi-complementarity between *some* peptide template pairs, see Figure 1(b). While essentially exponential growth of nucleic acids leads to constant proportions of the two complementary templates, in our simulations we observe almost exactly equal concentrations in the parabolic growth regime, despite the fact the replicative rate constants differ by even as much as an order of magnitude. This arises because the template-heterodimers act as a buffering population, a sort of chemical sink into which newly-formed quasi-complementary templates fall in pairs.

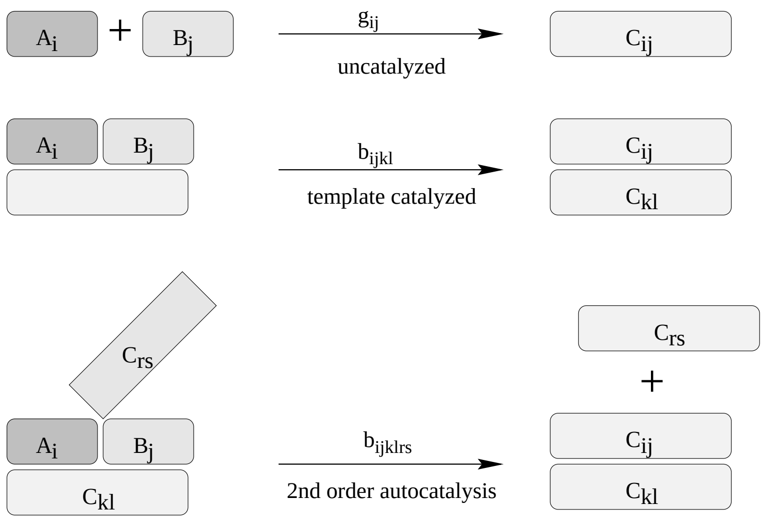

Figure 3. Modes of template formation. In complex systems of mixed templates, different modes of dynamic behavior are possible depending on the underlying mechanism of template synthesis. Uncatalyzed synthesis generally corresponds to linear growth. Template-instructed synthesis gives exponential or parabolic growth. The coupling of systems involving second order autocatalysis can give rise to hyperbolic growth, as has been studied for hypercycles [7].

The reaction system (4.1) which we have studied is a relative of the quasi-species model rather than the hypercycle [7]. In Figure 3 we illustrate the essential difference between systems that replicate in complementary or quasi-complementary fashion, and those that can be linked together to form classical hypercycles as they were defined by Eigen and Schuster [5, 7]. A realistic chemical system corresponding to the second order replicator equations, and in particular to a hypercycle, is represented by the scheme

$$A_i + B_j + C_{kl} + C_{rs} \underset{\bar{a}_{ijkl}}{\overset{a_{ijkl}}{\rightleftharpoons}} A_iB_jC_{kl} + C_{rs} \overset{b_{ijklrs}}{\longrightarrow} C_{ij}C_{kl} + C_{rs} \underset{\bar{d}_{ijkl}}{\overset{d_{ijkl}}{\rightleftharpoons}} C_{ij} + C_{kl} + C_{rs} \tag{8.1}$$

Here the template C_{rs} acts as an active ligase for the template-directed replication step $A_iB_jC_{kl} + C_{rs} \overset{b_{ijklrs}}{\longrightarrow} C_{ij}C_{kl} + C_{rs}$. Cyclical linking of such replication steps produces hypercycles. More generally, these systems can display an enormous diversity of dynamic behavior [12].

First because of their simplicity and second because of their generality, systems described by the scheme (4.1) are likely to feature prominently in both theoretical and experimental studies of the origins of life. As a modest goal we hope to develop kinetic signatures of systems involving many mutually replicating templates, and thus assist in the interpretation of experimental results without requiring the complete deconvolution of the individual growth curves.

Inclusion of the allosteric activation and inhibition of ligation steps will modify the reaction scheme down the lines of (8.1), and may be expected to produce collectively autocatalytic systems with multiple dynamic attractors, including limit cycles and chaotic attractors. Of special interest will be the minimal conditions under which coexistence and strong cooperation are observed within sets of templates of increasing diversity.

All of the systems in which molecular selection and collective autocatalysis have been studied have been purely exergonic from the thermodynamic point of view. Extensions to systems that are more realistically biological, linking together both exergonic and endergonic reactions and allowing the performance of thermodynamic work cycles to serve as the basis of autocatalytic reproduction, will open up new avenues of investigation [16].

Acknowledgments. This work was supported in part by NASA-AMES grant no. NAG 2-1091 and the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, Proj. No. P10857-CHE. The authors are grateful to the anonymous referees who commented on the manuscript, especially in relation to Figure 2.

Appendix A: Mean Fitness Is a Lyapunov Function in Szathmáry's Equation

Szathmáry's [28] model without spontaneous production is

$$\dot{x}_k = a_k x_k^p - x_k \sum_j a_j x_j^p \qquad a_k > 0, \quad 0 < p < 1\,. \tag{A.1}$$

Szathmáry showed that x_k increases when small. The *mean fitness* in this system is $\Phi(x) = \sum_j a_j x_j^p$.

Here we show a much stronger result† providing a complete analysis of the dynamical system (A.1).

Theorem. The rescaled mean fitness $L(x) = \Phi(x) - \Phi(\hat{x})$, where $\hat{x}$ denotes the unique interior equilibrium of (A.1), is a strict Lyapunov function in the interior of the state space.

Proof. The additive constant $\Phi(\hat{x})$ enforces $L(\hat{x}) = 0$. We shall show below that $\Phi(x(t))$ is strictly increasing along all trajectories in the interior of state space. Instead of verifying explicitly that $\Phi(x)$, and hence $L(x)$, has its unique maximum x we may also use this and the following two known facts about (A.1) to establish immediately that L is strict Lyapunov function: (i) $\hat{x}$ is the only fixed point of (A.1) in the interior, i.e., $\hat{x}$ is the only saddle point of Φ in the interior, and (ii) $\hat{x}$ is a (locally) stable equilibrium, i.e., $\hat{x}$ is a local maximum of $\Phi(x)$.
A short computation shows

$$\dot{\Phi}(x) = \sum_j a_j p x_j^{p-1} \dot{x}_j = p \sum_j a_j^2 x_j^{2p-1} - p \sum_j a_j x_j^p \sum_j a_j x_j^p\,. \tag{A.2}$$

Set $s_j := a_j^2 x_j^{2p-1}$. Then

$$\dot{\Phi}/p = \sum_j s_j \sum_j x_j - \left(\sum_j \sqrt{s_j x_j} \right)^2 . \tag{A.3}$$

† While this manuscript was under review a paper by Varga and Szathmáry (1997) was published with a proof of the same result. Here we give a different and substantially shorter proof.

We claim that the r.h.s. of this equation is positive for all non-negative vectors $\vec{x} \neq \vec{o}$ and $\vec{s} \neq \vec{o}$. Set $b_j = x_j / \sum x_k$ and $c_j = s_j / \sum s_k$. Dividing the r.h.s. by $\sum_j s_j \sum_j x_j$ we obtain

$$1 - \left(\sum_j \sqrt{(s_j / \sum_k s_k)(x_j / \sum_k x_k)} \right)^2 = 1 - \left(\sum_j \sqrt{b_j c_j} \right)^2 \geq 1 - \left(\sum_j \frac{b_j + c_j}{2} \right)^2 = 1 - 1 = 0\,. \qquad (A.4)$$

Here we have used the inequality between arithmetic and geometric mean. Hence $\dot{\Phi}(x) \geq 0$ if $x_k > 0$, i.e., the mean fitness is a Lyapunov function.

Equality hold if and only if both means agree for all j, i.e., if $b_j = c_j$ for all j. Returning to the original variables $b_j = c_j$ translates to

$$\hat{x}_j = s_j / \sum_i s_i = a_j^2 \hat{x}_j^{2p-1} / \sum_i s_i \quad \Longrightarrow \quad \hat{x}_j = \sqrt[1-p]{a_j} / \sum_i s_i = \sqrt[1-p]{a_j} / \sum_i \sqrt[1-p]{a_i}\,. \qquad (A.5)$$

For the last equality we have used that $\sum_j x_j = 1$.

q.e.d.

Therefore, every trajectory with an initial condition in the interior of the space approaches $\hat{x}$, and we may state the

Corollary 1. The interior fixed point $\hat{x}$ is globally stable in the interior of state space. An immediate consequence is

Corollary 2. The dynamical system (A.1) is permanent.

Since the dynamics on the boundary faces (on which one or more species vanish) is again of the form (A.1), only with a restricted set of variables, the theorem is valid there as well, i.e., there is a fixed point in the interior of each face of the state space which is globally stable within the face but unstable with respect to the invasion of any additional species. The theorem therefore provides complete knowledge of the qualitative behavior of equ.(A.1).

Appendix B: A Simple Numerical Example

Assume correct instruction and that the hetero-complex are unstable. We may then relabel the template species with a single index ($kl \to k$ and $ij \to i$). Equ.(7.1) reduces to

$$c_k = a_k[C_k] + b_k[C_k]^2\,, \qquad (B.1)$$

where a_k and b_k are complicated expressions of the physical rate constants. For our purpose here it is only necessary to note that a_k and b_k are independent parameters.

Using relative concentrations $x_k = c_k/c$ and $y_k = [C_k]/c$ this becomes

$$x_k = a_k y_k + b_k c y_k^2\,. \qquad (B.2)$$

Solving for y_k yields

$$y_k = \alpha_k(\sqrt{\beta_k x_k + 1} - 1) = \alpha_k \beta_k \varphi(\beta_k x_k) \qquad (B.3)$$

with the function $\varphi(z) = (\sqrt{z+1} - 1)/z$. It is easy to verify that $\varphi(z) \to 1/2$ for $z \to 0$. The differential equation (7.2) therefore becomes a replicator equation of the

$$\dot{x}_k = x_k\big(\alpha_k \beta_k \varphi(\beta_k x_k) - \Phi(x)\big) \qquad (B.4)$$

where $\Phi(x)$ is the mean fitness defined in Appendix A. Note the for large value of β_k we have effectively $\varphi(z) \sim 1/\sqrt{z}$ (and hence parabolic growth), while for small β_k we have $\varphi(z) \sim 1/2$ (and hence exponential growth). A detailed analysis of the replicator equation (B.4) will be presented elsewhere (Stadler, Stadler & Wills, in preparation).

In Figure 4 we show that depending on the parameters α_k and β_k we may find survival of single fittest species, the global coexistence that is characteristic of parabolic growth, or the extinction of some species while others coexist.

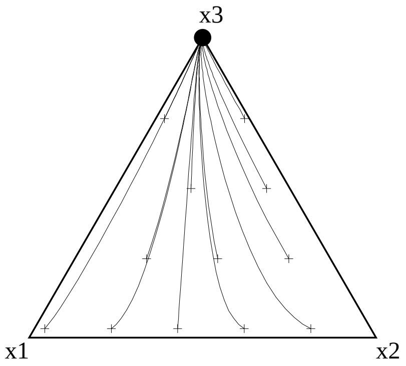

Figure 4. Examples of phase portraits for the evolution of systems governed by Equ.(B.4). In all cases $\alpha_1 = \alpha_2 = \alpha_3 = 1.0$. The parameter choice $\beta_1 = \beta_2 = 1.0, \beta_3 = 1.3$ leads to exclusive selection of template 3, $\beta_1 = 1.0, \beta_2 = 0.97$ and $\beta_3 = 1.02$ leads to universal coexistence and $\beta_1 = 1.19, \beta_2 = 1.23, \beta_3 = 1.0$ leads to coexistence of templates 1 and 2 and exclusion of template 3.

References

[1] T. Achilles and G. von Kiedrowski. A self-replicating system from three starting materials. *J. Angew. Chem. Int. Edn. engl.*, 32:1198–1201, 1993.

[2] C. Biebricher, M. Eigen, and W. C. Gardiner, Jr. Quantitative analysis of selection and mutation in self-replicating RNA. In L. Peliti, editor, *Biologically Inspired Physics*, pages 317–337, New York, 1991. Plenum Press.

[3] C. K. Biebricher, M. Eigen, W. C. Gardiner Jr., and R. Luce. Kinetics of RNA replication: Plus-minus asymmetry and double strand formation. *Biochemistry*, 24:3186–3194, 1985.

[4] 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.

[5] M. Eigen. Selforganization of matter and the evolution of macromolecules. *Naturwiss.*, 58:465–523, 1971.

[6] M. Eigen, J. S. McCaskill, and P. Schuster. The molecular quasi-species. *Adv. Chem. Phys.*, 75:149–263, 1989.

[7] M. Eigen and P. Schuster. *The Hypercycle.* Springer-Verlag, New York, Berlin, 1979.

[8] M. Famulok, J. Nowick, and J. Rebek Jr. Self-Replicating Systems. *Act.Chim.Scand.*, 46(4):315–324, 1992.

[9] R. Hecht, R. Happel, P. Schuster, and P. F. Stadler. Autocatalytic networks with intermediates I: Irreversible reactions. *Math. Biosc.*, 140:33–74, 1997.

[10] M. W. Hirsch and S. Smale. *Differential Equations, Dynamical Systems and Linear Algebra.* Academic Press, New York, 1974.

[11] J. Hofbauer. The selection mutation equation. *J. Math. Biol.*, 23:41–53, 1985.

[12] J. Hofbauer and K. Sigmund. *Dynamical Systems and the Theory of Evolution.* Cambridge University Press, Cambridge U.K., 1988.

[13] V. Hutson and K. Schmitt. Permanence and the dynamics of biological systems. *Math. Biosc.*, 111:1–71, 1992.

[14] K. D. James and A. D. Ellington. The search for missing links between self-replicating nucleic acids and the RNA world. *Origins Life Evol. Biosphere*, 25:515–530, 1995.

[15] K. D. James and A. D. Ellington. Catalysis in the RNA world. In A. Brack, editor, *The Molecular Origins of Life*, page in press. Cambridge University Press, Cambridge UK, 1997.

[16] S. A. Kauffman. Investigations. Technical report, Santa Fe Institute, Santa Fe, NM, 1996. #96-08-072.

[17] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, and M. R. Ghadiri. A self-replicating peptide. *Nature*, 382:525–528, 1996.

[18] 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.

[19] L. Michaelis and M. I. Menton. Die Kinetic der Invertinwirkung. *Biochem. Z.*, 49:333–369, 1913.

[20] V. Rotello, J.-I. Hong, and J. Rebek Jr. Sigmoidal growth in a self-replicating system. *J. Amer. Chem. Soc.*, 113:9422–9424, 1991.

[21] P. Schuster, K. Sigmund, and R. Wolff. Dynamical systems under constant organization III: Cooperative and competitive behaviour of hypercycles. *J. Diff. Eqns.*, 32:357–368, 1979.

[22] L. A. Segel and M. Slemrod. The quasi-steady state assumption: a case study in perturbation. *SIAM Rev.*, 31:446–477, 1989.

[23] D. Sievers and G. von Kiedrowski. Self-replication of complementary nucleotide-based oligomers. *Nature*, 369:221–224, 1994.

[24] B. M. R. Stadler and P. F. Stadler. Dynamics of small autocatalytic reaction networks III: Monotonous growth functions. *Bull. Math. Biol.*, 53:469–485, 1991.

[25] P. F. Stadler. Complementary replication. *Math.Biosc.*, 107:83–109, 1991.

[26] P. F. Stadler, W. Schnabl, C. V. Forst, and P. Schuster. Dynamics of autocatalytic reaction networks II: Analytically treatable special cases. *Bull. Math. Biol.*, 57:21–61, 1995.

[27] P. F. Stadler and P. Schuster. Mutation in autocatalytic networks - an analysis based on perturbation theory. *J. Math. Biol.*, 30:597–631, 1992.

[28] E. Szathmáry and I. Gladkih. Sub-exponential growth and coexistence of non-enzymatically replicating templates. *J. Theor. Biol.*, 138:55–58, 1989.

[29] A. Terfort and G. von Kiedrowski. Self-replication by condensation of 3-aminobenzamidines and 2-formylphenoxyacetic acids. *Angew. Chem. Int. Ed. engl.*, 31:654–665, 1992.

[30] S. Varga and E. Szathmáry. An extremum principle for parabolic competition. *Bull. Math. Biol.*, 59:1145–1154, 1997.

[31] G. von Kiedrowski. A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed. Engl.*, 25:932–935, 1986.

[32] 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.

[33] 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.

[34] W. S. Zielinski and L. E. Orgel. Autocatalytic synthesis of a tetranucleotide analogue. *Nature*, 327:346–347, 1987.