Eur. Phys. J. B **38**, 103–110 (2004)
DOI: 10.1140/epjb/e2004-00105-2

THE EUROPEAN PHYSICAL JOURNAL B

Evolutionary stabilization of generous replicases by complex formation

R.M. Füchslin[a], S. Altmeyer, and J.S. McCaskill

Fraunhofer – Biomolecular Information Processing Schloss Birlinghoven, St. Augustin 53754, Germany

Received 17 December 2003
Published online 20 April 2004 –

Abstract. The importance of spatial organization for the evolutionary stability of trans-acting replicase systems ($T \xrightarrow{X} 2T$), where T is a general member of a combinatorial family including the special catalyst X, is now well established, by analytical and Monte Carlo models [1,2]. Complex formation as an intermediate step in replication ($X + T \rightleftharpoons XT \rightarrow X + 2T$), besides refining the model, enhances co-localization of replicase X and templates T and is shown here to thereby contribute to the evolutionary stability of the catalyst. Applying the established individual molecule stochastic PRESS-framework [3], the performances of cooperative replication with and without intermediate complex formation are compared, and the beneficial effect of complex formation as enhancing stability is studied numerically under various conditions. The results obtained are of value for studies of prebiotic evolution, but also point towards a possible mechanism for stabilizing replication systems in adaptive molecular engineering.

PACS. 87.16.Ac Theory and modeling; computer simulation – 87.23.Cc Population dynamics and ecological pattern formation – 87.23.Kg Dynamics of evolution

1 Introduction

The evolutionary generation of biomolecular information is based on interacting molecular replication processes acting in the presence of physical constraints such as molecular loss (outflow, decomposition) and mutation events. These constraints, which must have dominated in early evolution, but which have become highly channeled in modern cells, have now again become important in attempts to evolve catalytic function in vitro in the laboratory. Although much work has been done on combinatorial families of single molecule replicator systems of the form $X \rightarrow 2X$ [4–6] (such systems can be studied in the laboratory using modern polymerases as a fixed component of the environment [7]), autonomous evolution generally must involve generically trans-acting catalysts (such as generic polymerases) and these are the most interesting targets for evolutionary biotechnology. The overall kinetics is then non-linear, as in the hypercycle [8] kinetics, having the simplified stoichiometry $2X \rightarrow 3X$.

Studies of both prebiotic evolution and molecular engineering of primitive pre-cellular replication systems naturally split into considerations about the chemical nature of the constituents of such a system (RNA, proteins, etc.) and questions concerning the principal possibility for the stabilization of different kinetic mechanisms. This paper deals with the latter issue. Since biochemical catalysts usually do not act directly on themselves but on other molecules of a specific type, trans-acting (also called "cooperative") replication, modeled as a single step, and characterized by the reaction

$$T \xrightarrow{X} 2T \quad \text{or} \quad X + T \rightarrow X + 2T, \tag{1}$$

is the generic form of catalyzed self-replication. An evolvable biochemical realization of such a system may consist of a combinatorially large family of chain molecules with general member T which contains a small subset of catalysts X capable of catalyzing the replication of any other chain molecule out of the whole family (note that Eq. (1) includes the special case of the general template T being also a catalyst X).

Biochemical replication processes are generically error prone. Starting with a pure population of catalysts X, a replication system has to be able to cope with occasional exploiting mutants Y (sometimes termed parasites), i.e. molecules which can be replicated by X, thereby using resources, but which do not contribute to the replication process themselves. Such mutations have been shown [9–11] to doom a cooperative replication system to extinction in the absence of an ancillary stabilization mechanism. Although the efficiency of spatial effects

[a] e-mail: rudolf.fuechslin@biomip.fhg.de

for stabilizing sequence information was first proposed [1] for the hierarchical extension of the quasispecies theory to hypercyclic coupled systems of self-replicators [8], in spatially resolved stochastic kinetics, simple alternatives to hypercyclic coupling have been proposed [2,12] for systems in which a non-specific replicase species copies all others independently of sequence. In all cases, the necessary correlations between genetic sequences encoding the replicase function and their preferential proliferation by replicases are achieved spatially, in the simplest case merely by the effects of a finite diffusion constant.

A basic evolvable trans-catalytic replication system (in the following referred to as the generous replicator model, or XY-model for short), which can be stabilized by spatial organization, is given by

$$\begin{aligned} 2X &\xrightarrow{r_B Q} 3X, \\ 2X &\xrightarrow{r_B(1-Q)} 2X + Y, \\ X + Y &\xrightarrow{r_B r} X + 2Y, \\ X, Y &\xrightarrow{d} \varnothing. \end{aligned} \tag{2}$$

The total rate of catalyzed synthesis by the replicase X is determined by the rate parameter r_B. The replication fidelity is given by the quality factor Q, r denotes a possible parasite advantage and a uniform rate d models the decay or outflow of sequences. The decay rate d is assumed to be uniform for both a theoretical and an experimental reason; the set of parameters should be as small as possible and second, non-specific outflows are also easy to implement in an experiment by dilution. The term "generous replicator" is explained by the fact that, as long as $r \geq 1$, there is no exploiter marginalization whatsoever.

With explicit resource consumption during replication (e.g. monomers), this reaction has been shown both by deterministic kinetic [12] modeling (with resource limitation and a distinct higher diffusion coefficients for the resource) and by stochastic modeling, even with equal diffusion coefficients [13], to create self-replicating spot patterns which are favourable to the evolutionary stabilization of the cooperative replicase function. Furthermore, recent theoretical progress allowed the authors to study the spatially structured dynamics of the trans-catalytic generous replicator system equation (2), with empty site modeling of resource constraints, on a completely interconnected topology (simplex) even analytically [3], and to show that the results carry over to describe essential evolutionary behaviour in physical space. This makes the generous replicator model a reliable starting point for the study of the effects of more complex (and more natural) reaction schemes.

Since the information dynamics in the evolving generous replicator system can be characterized as a struggle to maintain local correlations between a functional (polymerase) catalyst and the informational molecules which encode it, it is natural to ask whether complex formation between the polymerase and its target template cannot provide a major enhancement of these positive correlations through joint diffusion, and so stabilize trans-acting systems against evolutionary exploitation. Complex formation is in fact a physically necessary intermediate step in trans-catalysis, as in classical Michaelis-Menten kinetics (though, as the authors recognize, still not reflecting the full complexity of the process). The corresponding system to equation (2), complex forming generous replicators,

$$\begin{aligned} 2X &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C_{XX} \xrightarrow{k_2 Q} 3X, \\ 2X &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C_{XX} \xrightarrow{k_2(1-Q)} 2X + Y, \\ X + Y &\underset{k_{-1}}{\overset{rk_1}{\rightleftharpoons}} C_{XY} \xrightarrow{k_2 r} X + 2Y, \\ X, Y, C_{XX}, C_{XY} &\xrightarrow{d} \varnothing, \end{aligned} \tag{3}$$

in the following also referred to as the CXY-model, is the subject of this paper. It will be seen to become a more significant correction as the informational molecules become longer under the influence of evolution. Note that the kinetics of a model with complex formation can be smoothly mapped onto the kinetics of the pure generous replicator model by tuning the rate parameters k_i towards a situation where $k_1 \ll k_{-1} \ll k_2$.

In order to understand the key issue here more fully, it is fruitful to precede the quantitative analysis below with a discussion of the way in which complex formation in the generous replicator model can contribute to evolutionary stability. Oversimplified, the stabilizing effect of spatial organization is caused by the fact that in regions infected by exploiters (cf. defective interfering parasites [14]), the replication process will eventually break down through loss of catalysts and the entire region will then be depleted, a fate which is not shared by clusters containing only or mostly correct sequences [15]. The latter chemically active clusters may then spawn into void areas and seed new clusters, whereby it is crucial those molecules with sequences leading to replicase functionality travel from regions with high chemical activity to void areas and meet there. The chance of meeting another molecule with correct sequence in such sparsely occupied areas is small. If, however, the intact catalyst and template molecules stay bound for long enough to co-migrate, the chance to reach a void area and initiate a new population is enhanced.

The main goal of this work is to show that the effect of complex formation is not only a byproduct of a more complicated replication procedure but by itself a highly significant stabilizing mechanism. This is achieved by determining the parameter region in which the generous replicator model can be stabilized and showing that the inclusion of complex formation allows stabilization outside this region for kinetically equivalent or even less favorable parameter values of the complex forming generous replicator model. Evolution will naturally make use of this mechanism to build up catalytic information.

The article is organized as follows: in Section 2, the PRESS-method, a more efficient procedure than Monte-Carlo simulation which allows one to study the stochastic dynamics of evolvable systems with spatially structured molecular distributions, is presented in condensed form and applied to the CXY-model. Section 3 presents evidence for the claim that complex formation indeed gives rise to additional stability of a replication system. In Section 4, some alternatives to the complex formation mechanism and some implications of the results obtained are discussed.

2 Methods – the PRESS framework

2.1 The basic formalism

Spatial structure and stochasticity have already been recognized as physical mechanisms, which can stabilize a self-encoding catalytic replication system against exploitation, by proliferating non-coding genetic material. Such processes are difficult to simulate over long evolutionary time scales, simply because of their computational demand (calculations are possible on special purpose machines [15], but a full spatially and sequentially resolved simulation is not feasible on a conventional workstation). Furthermore, single simulations do not give analytical insight into the interplay of different system parameters. However, for suitable classifications of sequences, it is possible to set up an analytical and integrable treatment of individual-based probability dynamics, the PRESS-framework [3] (Probability Reduced Evolution of Spatially discrete Species), which models qualitatively and even quantitatively well the dynamics of a replication system and can be implemented efficiently on conventional hardware. The PRESS-framework is discussed in full detail in [3]. Here the method is applied to the complex formation kinetics, after a brief summary of the pure generous replicator model for reference.

The basic simplification of the PRESS-framework is the assumption that the overall stability of a replication system distributed on different sites is accounted for at the individual molecule level by local probabilistic dynamics being coupled (via diffusion) to an ensemble of other statistically equivalent local population sites. The physical topology for which this self-consistent probability-field approach, pursued in the PRESS-framework, is exact is that of an infinite number of completely interconnected sites of limited size (i.e. an infinite dimensional simplex). The individual sites are taken to be small enough to be well mixed on the timescale of reactions.

2.2 The generous replicator model in the PRESS-framework [3]

In the case of the generous replicator model, each site can be described by a state (X, Y), where X stands for the number of wild type molecules on the site and Y represents the number of catalytically defective molecules. We use capital letters to distinguish these stochastic variables from lower case deterministic concentrations, and trust that no confusion will arise with the use of upper case symbols to describe the molecular types themselves. For all sites, $X+Y \leq N$, where N is the maximum site population size (carrying capacity). Consequently, for S distinguished classes of molecules (in case of the generous replicator model $S = 2$) there are $\binom{N+S}{S}$ different states. The Markovian dynamics is characterized by probabilistic transitions between different states, reflecting either reactions or molecular exchange between sites. Assuming an infinite number of mutually connected sites (all sites "see" the same environment), there are no correlations with the states of specific neighbor sites. This then means that the state of the system can be described by a single discrete probability distribution function $P(X, Y; t)$ (later collected in a probability vector $\vec{P}(t)$ with an appropriate numbering of the states (X, Y)). The time evolution of this probability distribution can be formulated as a set of coupled ODEs that are derived from the transition rates w_{ij} between different states. As in linear master equations, the transition matrix L, determines the dynamics of $\vec{P}(t)$ according to

$$\frac{d\vec{P}}{dt} = L\vec{P}, \quad \text{where} \quad L_{ij} = w_{ij} - \delta_{ij} \sum_k w_{kj}. \tag{4}$$

The structure of the diffusion-induced transitions leads to a non-linearity in this equation. The exchange of a X molecule on a site with a Y molecule from another site, for example, is given by

$$\begin{aligned} w^{diff.}_{X\to X-1, Y\to Y+1} &= D_\infty \frac{X}{N} \left(\sum_{X',Y'} Y' P(X', Y') \right), \\ &= D_\infty \frac{X}{N} \overline{Y} \end{aligned} \tag{5}$$

where D_∞ stands for the diffusion coefficient (in this high dimensional space) and $\overline{Y}$ for the average value of Y over the different sites. That the effect of diffusion is described in terms of the single site population probabilities themselves (which also determine the overall site averaged values of the different molecule types) is a key advantage of the highly symmetric topology used in the PRESS-framework. The result is a closed but quadratically non-linear master equation and the effect of co-localization becomes apparent by the fact that in general $P(X, Y) = P(X)P(Y)$ does not hold as would be the case in a well stirred system, or in the limit $D_\infty = \infty$.

Setting $Z = N - X - Y$, the complete set of transition probabilities for the generous replicator model is:

$$\begin{aligned}
w_{X\to X+1,Y\to Y} &= r_B Q X \frac{X-1}{N-1}\frac{Z}{N-2} + D_\infty \overline{X}\frac{Z}{N},\\
w_{X\to X,Y\to Y+1} &= r_B(1-Q)X\frac{X-1}{N-1}\frac{Z}{N-2}\\
&\quad + r r_B X \frac{Y}{N-1}\frac{Z}{N-2} + D_\infty \overline{Y}\frac{Z}{N},\\
w_{X\to X-1,Y\to Y} &= dX + D_\infty \overline{Z}\frac{X}{N},\\
w_{X\to X,Y\to Y-1} &= dY + D_\infty \overline{Z}\frac{Y}{N},\\
w_{X\to X+1,Y\to Y-1} &= D_\infty \overline{X}\frac{Y}{N},\\
w_{X\to X-1,Y\to Y+1} &= D_\infty \overline{Y}\frac{X}{N}
\end{aligned} \tag{6}$$

with parameters as defined in equation (2). Importantly, the reactions take into account local limitations by resources, e.g. replication is formulated according to $X + X + \varnothing \to X + X + X$ with $\varnothing$ representing a consumable, limited resource. We do not model the resource replenishment process explicitly, although this could be done, but assume that resource repletion chemistry is kinetically linked to the destruction processes. This is consistent with the focus of the work being on the information dynamics rather than a particular choice of underlying chemistry. With this in mind, the denominators $(N-1)$ and $(N-2)$ properly reflect the effect of the finite chemical capacity of the sites.

For the generous replicator model, the PRESS-framework not only models the qualitative and quantitative behavior of a full spatially resolved setting well but also allows several analytical results to be gained. In the case of small N, even a full analytical solution for the stationary states of the system was obtained, which allowed the authors to calculate critical bounds for the system parameters and to get insight into the way they affect the system.

In summary, the PRESS-framework provides a mathematical framework that allows the effects of heterogeneity and stochasticity to be studied, without having to care about (and to compute) the subtle pattern forming mechanisms (building up regions of different composition) in a finite dimensional environment. Projection of the results of the simplex topology onto finite dimensional space, using a renormalization of diffusion coefficients [3], proved successful, so that we expect the influence of specific patterns in the evolutionary stability to be of secondary importance.

2.3 Complex formation in the PRESS-framework

The state of a site in the complex forming generous replicator model introduced above is given by the discrete random variables (X, Y, C_{XX}, C_{XY}), determining the numbers of molecules or complexes of each kind. The transition rates are derived basically in the same way as in the pure generous replicator model. In order to avoid lengthy notation, only the chemical transitions are given. Diffusive processes are formulated according to equation (5), but taking into account the enlarged set of molecules.

$$\begin{aligned}
w_{X\to X-2,Y\to Y,C_{XX}\to C_{XX}+1,C_{XY}\to C_{XY}} &= k_1 X \frac{(X-1)}{(N-1)},\\
w_{X\to X-1,Y\to Y-1,C_{XX}\to C_{XX},C_{XY}\to C_{XY}+1} &= \\
&\quad r k_1 X \frac{Y}{(N-1)},\\
w_{X\to X+2,Y\to Y,C_{XX}\to C_{XX}-1,C_{XY}\to C_{XY}} &= \\
&\quad k_{-1} C_{XX}\frac{Z}{(N-1)},\\
w_{X\to X+1,Y\to Y+1,C_{XX}\to C_{XX},C_{XY}\to C_{XY}-1} &= \\
&\quad k_{-1} C_{XY}\frac{Z}{(N-1)},\\
w_{X\to X+3,Y\to Y,C_{XX}\to C_{XX}-1,C_{XY}\to C_{XY}} &= \\
&\quad k_2 Q C_{XX}\frac{Z(Z-1)}{(N-1)(N-2)},\\
w_{X\to X+2,Y\to Y+1,C_{XX}\to C_{XX}-1,C_{XY}\to C_{XY}} &= \\
&\quad k_2(1-Q) C_{XX}\frac{Z(Z-1)}{(N-1)(N-2)},\\
w_{X\to X+1,Y\to Y+2,C_{XX}\to C_{XX},C_{XY}\to C_{XY}-1} &= \\
&\quad r k_2 C_{XY}\frac{Z(Z-1)}{(N-1)(N-2)},\\
w_{X\to X-1,Y\to Y,C_{XX}\to C_{XX},C_{XY}\to C_{XY}} &= dX,\\
w_{X\to X1,Y\to Y-1,C_{XX}\to C_{XX},C_{XY}\to C_{XY}} &= dY,\\
w_{X\to X,Y\to Y,C_{XX}\to C_{XX}-1,C_{XY}\to C_{XY}} &= dC_{XX},\\
w_{X\to X,Y\to Y,C_{XX}\to C_{XX},C_{XY}\to C_{XY}-1} &= dC_{XY}.
\end{aligned} \tag{7}$$

Here, Z stands for $N - X - Y - C_{XX} - C_{XY}$, i.e. the number of unconsumed resource tokens on a side. Note that the rate of complex formation is twice as high for two molecules of type X on an otherwise empty site compared to a situation with one X and one Y. This would be wrong if complex formation is a completely symmetric process, but is correct when, as expected, there are two asymmetric binding sites on a molecule of type X, one as a catalyst and one as a template. In the case of a C_{XX}-complex, the question which of the two molecules of type X acts as catalyst and which as a template may depend on the details of the collision of the two molecules.

3 Results

3.1 Shift and enlargement of the range of possible diffusion rates of the complex forming generous replicator model in a simplex topology

In order to characterize the effects of complex formation, the parameter ranges leading to stabilization for

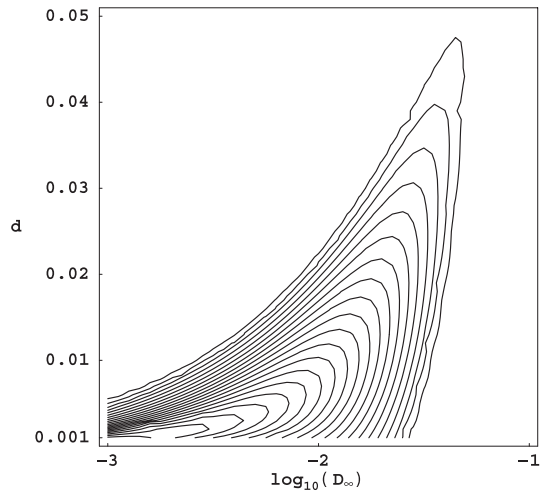

Fig. 1. Contour plot of the critical error rate R_C of the basic XY-model as calculated within the PRESS-framework on a simplex with sites of size 3 as a function of the migration rate D_∞ and the decay rate d. Since the time has been rescaled to give $r_B = 1$, all rates are given with respect to r_B. For a given d, stabilization ($R_C \neq 0$) is only possible in a limited range of D_∞. The maximal R_C obtained in this figure equals 0.056.

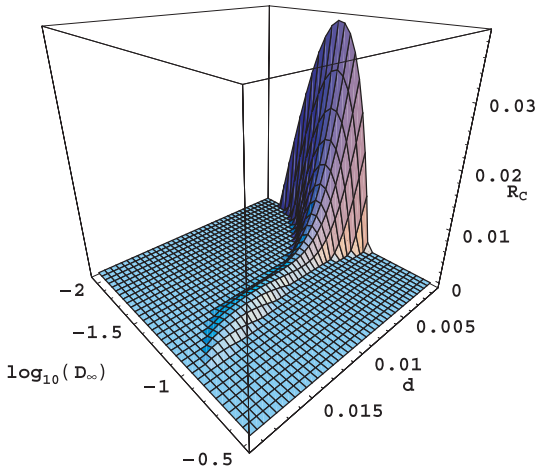

Fig. 2. Three-dimensional overview of the region of stabilization in the (D_∞, d) – plane for the CXY-model under the same conditions as described in Figure 1 and with parameters $k_1 = 0.085$, $k_{-1} = 0.001$, $k_2 = 1$, $r = 1.01$. Note that stabilization is also possible for diffusion rates which are not attainable by the XY-model.

both models are compared with one another. In order to avoid lengthy expressions, we use in this section the previously defined abbreviations XY- and CXY-model. The XY-model is dependent on the kinetic rate parameters D_∞, d, r_B, the replication fidelity Q and the relative advantage of exploiting templates r. r_B can be set to one without loss of generality (rescaling of time). The exploiter advantage factor r is set to 1.01, following experience with the XY-model. This value represents a 1% advantage for exploiting templates, allowing a clearer separation of exploitation from the physical effects of errors than the otherwise more natural "neutral" choice, $r = 1$, for which the net exploiting advantage vanishes (slowly) as $Q \to 1$ (note that even with $r = 1$, the replicator is copied with a fidelity $Q < 1$, whereas the descendant of an exploiter is always again of the exploiting type).

The performance of the XY-model was characterized by numerically calculating the critical error rate $R_C = 1 - Q_C$ necessary for stabilizing the system as a function of the diffusion coefficient D_∞ and the decay rate d, as shown in Figure 1. Here, stabilization means that the system either reaches a stationary state or enters a limit cycle (for a more detailed study of the dynamical properties of the XY-model see [3]). It is emphasized that the limits of viability in the XY-model (with r set to 1.01) are completely described by the $R_C(d, D_\infty)$ function shown.

The CXY-model has three parameters, k_1, k_{-1} and k_2, describing the detailed kinetics of complex formation and replication, instead of the single parameter r_B in the XY-model. Time was rescaled so that the overall rate of replication was unity ($r_B = 1$) in the XY-model. In order to guarantee that the complex formation process in the CXY-model does not deliver any direct kinetic advantage, the first order replica release rate coefficient k_2 is set to unity and consequently the overall rate of replication is less than one. The alternative direct match of the overall rate of homogeneous replication is thwarted by the nonlinear concentration dependence of replication. We varied the remaining binding rate coefficients k_1 and k_{-1}. Figures 2 and 3 show the stabilization results for a favorable choice of these parameters ($k_1 = 0.085$, $k_{-1} = 0.001$) firstly in an overview and then for the particularly interesting region of higher diffusion rates. The value $k_1 = 0.085$ was found to be optimal in Figure 4, where R_C is plotted for fixed d and k_{-1} and variable k_1 and D_∞. Figure 5 shows a similar plot for the rate coefficient k_{-1}, which not so surprisingly shows that the lower k_{-1}, the better it is for system stability, with a maximum for $k_{-1} = 0$. Consequently, a finite but low value was adopted in the calculations for Figures 2 and 3.

More generally, the region of stabilization obtained differs importantly from that of the XY-model. For higher decay rates, stabilization is also possible in regions of the diffusion rate that are not attainable in the XY-model. It is emphasized that the exploiter advantage factor r enters equations (3) both in complex formation and in replication which means that the advantageous behavior of a complex formatting system can neither be explained directly by kinetic reasoning (see above) nor by any specific suppression of exploiter advantage.

3.2 The complex forming generous replicator model in physical space

The stabilizing effect of complex formation is not an artifact of the simplex topology as can be seen in Figure 6. Here we compare the behavior of the critical error rate R_C in finite dimensions to the results in the simplex case. It turns out that these values, calculated in a computationally demanding calculation on a rectangular grid, are

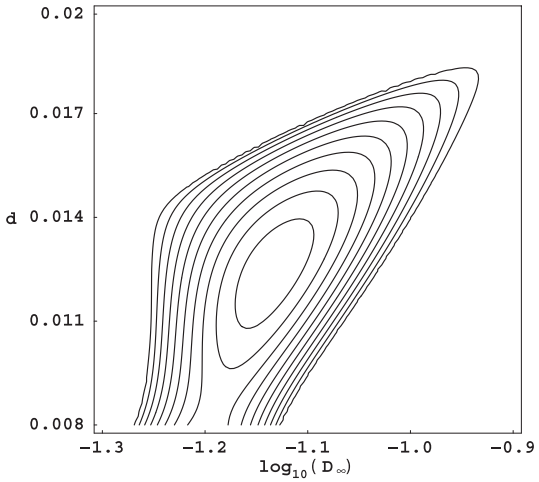

Fig. 3. Detailed contour plot of the region of stabilization accounting for higher migration rate in Figure 2 of the CXY-model. The maximal R_C obtained equals 0.0017.

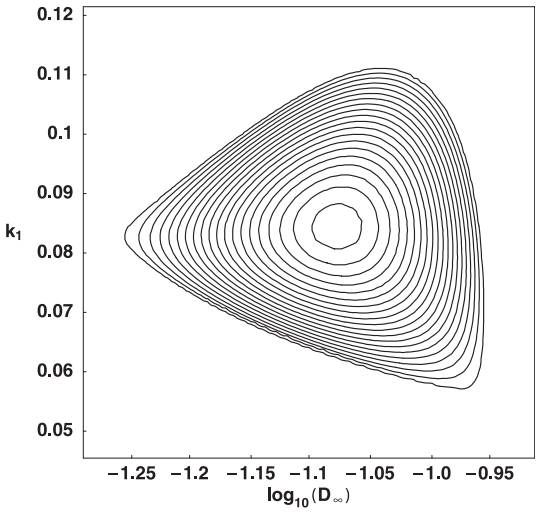

Fig. 4. Contour plot of R_C for the CXY-model in the (D_∞, k_1)–plane for parameters $d = 0.015, k_{-1} = 0.001, k_2 = 1, r = 1.01$. A maximal $R_C = 0.0015$ is obtained for $k_1 = 0.085$.

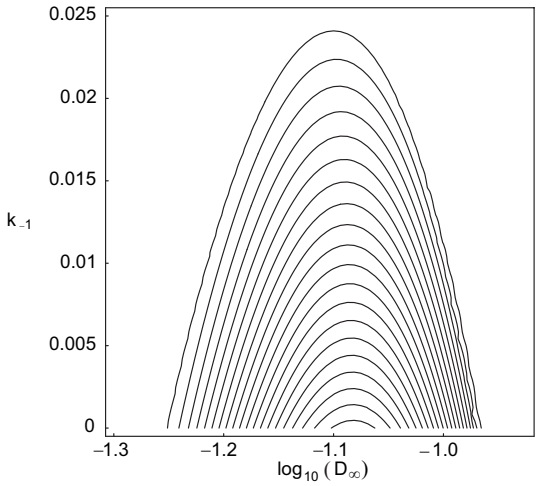

Fig. 5. Contour plot of R_C for the CXY-model in the (D_∞, k_{-1})–plane for parameters $d = 0.015$, $k_1 = 0.085$, $k_2 = 1$, $r = 1.01$. A maximal $R_C = 0.0017$. is obtained for $k_{-1} = 0$, which physically can only be reached in a limiting case.

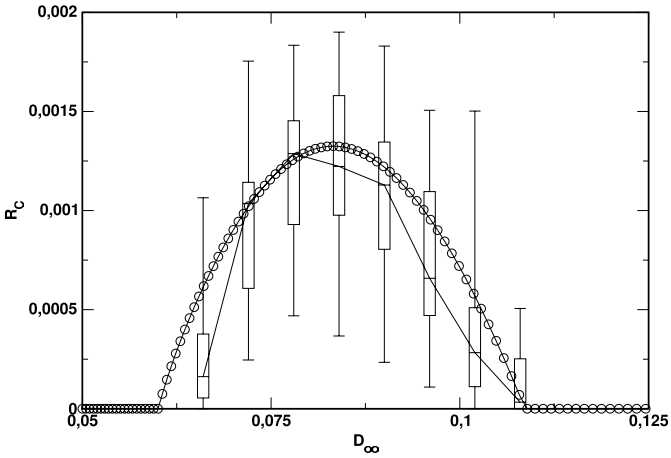

Fig. 6. Comparison of R_C, calculated on the infinite simplex (circles) with the corresponding results for stochastic simulations on a three dimensional grid (boxes). The parameters used are $d = 0.015$, $k_1 = 0.085$, $k_{-1} = 0.001$, $k_2 = 1$, $r = 1.01$. The diffusion parameter D_3 is rescaled according to equation (8). The stochastically evaluated R_C values show a considerable spread (as indicated by the quartile boxes, referring to at least 25 individual values) depending on the random initial configuration. The median values of the distributions for finite systems in three dimensions, however, are well reflected by the much easier to determine R_C values for the simplex.

predicted well by the much more efficient PRESS-method on the simplex. The results, however, have to be interpreted with care. Calculations in finite dimensions have generically to be performed with finite population sizes. But every finite population will eventually die out due to stochastic fluctuations. This in turn means that one has to invoke some cut-off criterion for stability, which is here defined as the ability to survive a sufficiently long time after the initial transient. The box plot in Figure 6 reflects the quartiles and the median of the resulting distribution of R_C values. These statistics where chosen because the fate of the system depends strongly on the randomly chosen initial configuration, even for a three dimensional grid of size $64 \times 64 \times 64$. It is apparent from the figure that the median of the R_C distribution is approximated well by the prediction of the PRESS-method, whereby parameters has to be rescaled. For simulations in an environment of dimension n (and assuming a rectangular grid of sites), a rescaling of the diffusion rate according to

$$D_n = \frac{2n-1}{2n} D_\infty \tag{8}$$

turned out to be appropriate (for a discussion, see [3]).

4 Discussion

The effect of a simplex topology on the evolution of altruistic functions was studied by Kimura [16]. However, Kimura postulated some group selection advantage of favorably composed sites, which is also a prerequisite of the stochastic corrector model [17,18]. The generous replicator model also showed evolutionary stabilization, but the advantage and emergence of spatial co-localization is not an assumption of the model but an intrinsic consequence of the system dynamics. This makes the evolutionary fate of a system sensitive to details of the kinetic mechanism and allows for non-trivial advantages of processes such as complex formation, which are not obvious at a first glance (as a matter of fact, complex formation slows down replication, since it transforms a one-step procedure into a two step reaction). It has been shown, even without the adaptation of the relevant kinetic parameters, that the effect of correlated transport allows the stabilization of the catalysts in otherwise inaccessible regions of the parameter space.

Cooperative replication cannot be studied by fitness landscapes (for an overview over different models based on fitness landscapes, see [19,20]), since the number of offspring of a molecule not only depends on its own nature, but also on the composition of its environment. Besides reflecting the fact that biochemical catalysis, especially replication, is generally a trans-acting process, cooperative replication offers a possibility to understand how sequences far from the wild type in sequence space can be reached. In contrast to investigations based on fitness landscapes, models of cooperative replication allow a kinetically unhindered replication of non-catalytic molecules (in the presented model, non-catalytic molecules can even be given a replication advantage) and therefore a rapid accumulation of mutations.

However, the PRESS-framework is based on some prerequisites which have to be discussed. Firstly, the site size is limited and needs to be. A site size of three has been chosen in the presented work because it is the minimal one possible. As a matter of fact, stabilization is not only possible in larger regions but even stronger (i.e. for higher R_C) for $N = 4$. However, for each fixed diffusion rate D_∞ there was in all calculations a maximal N above which stabilization was no longer possible. This limit is not surprising: as shown in Appendix A, a spatially homogeneous system cannot host a stationary replication system with complexes (although we have not extended the proof of this to include oscillatory solutions).

Secondly, the possibility of complexes of the form C_{YY} (which do not lead to replication but to correlated transport of non-catalytic molecules) was not included. This is because we did not find this process to be advantageous for the exploiting species, i.e. the evolution of such a binding to be a problem for the stability of the system. On the contrary, as can be seen in Figure 7, the formation of parasite complexes is beneficial for the wild type.

Thirdly, the possibility of complex formation does indirectly enhance the carrying capacity of the system. A possible way to avoid this would be to invoke a somewhat artificial constraint on the possible states of a site, namely $X + Y + 2C_{XX} + 2C_{XY} \leq N$, which means that complexes are counted as two molecules. The problem with this approach is that diffusion can no longer be modeled according to equation (5), since the probability for e.g. the exchange of a complex with an empty location $\varnothing$ does not depend on $\bar{\varnothing}$ but on the actual probability $P(X, Y, C_{XX}, C_{XY}; t)$ of states able to host an additional complex. This leads to complex, non-constant diffusion coefficients that depend on the system state. However, even under this restriction of the carrying capacity, stabilization in otherwise inaccessible regions of the parameter space still persists, as can be seen in Figure 8, calculated with different kinetic rate parameters and with a smaller exploiter advantage r.

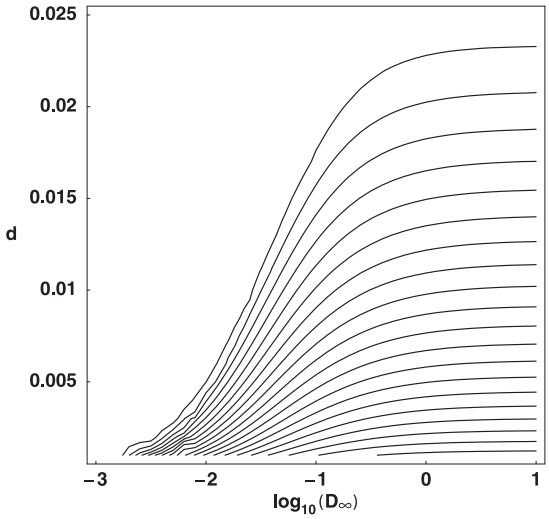

Fig. 7. Contour plot of R_C for the CXY-model in the (D_∞, d)–plane for parameters $k_1 = 0.085$, $k_2 = 1$, $r = 1.01$ under inclusion of possible formation of C_{YY}-complexes. A large maximal $R_C = 0.84$ is obtained. This result shows that the formation of exploiter complexes is not favorable for them and therefore evolutionarily avoided.

In summary, the calculations presented show that spatial organization does not only give rise to stabilization of more natural replication schemes than that of the pure generous replicator model, but also that the specific properties of complex formation allow stabilization in regions of the parameter space unattainable for the generous replicator model without complexes. Since the kinetic parameters are chosen in a manner that excludes an explanation of this effect by direct kinetic advantage, complex formation itself must interact with the spatial structure to provide an advantage for stabilization of replication systems.

Further, the basic simplification of the PRESS-method, namely its assumption of a simplex topology, turned out to yield large efficiency gains but to reflect the effect of co-localization in a manner that allows reliable estimates also for systems embedded in a finite dimensional environment.

The study of simple replication systems is not only relevant with respect to prebiotic evolution and the origin of

$$k_{-1} = -\frac{(k_2^2(1-Q)r\theta^4 x + d^2 y)(x+y) + dk_2\theta^2((1-Q)x^2 + (3r-Q-1)xy + y^2)}{\theta(dy(x+y) + k_2\theta^2 x((1-Q)x + (r-Q)y)}. \quad (10)$$

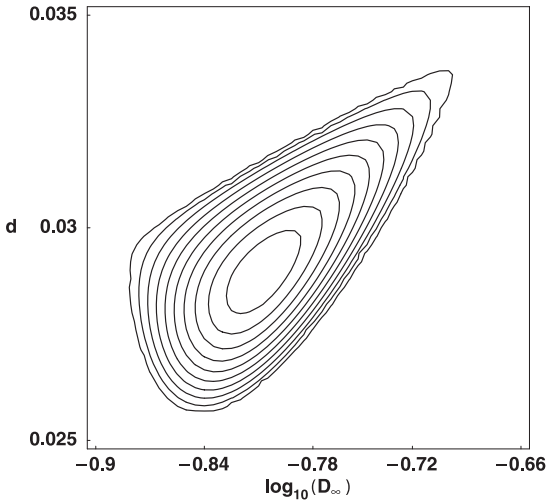

Fig. 8. Contour plot of R_C for the CXY-model in the (D_∞, d)–plane under the assumption of a carrying-capacity restriction according to $X + Y + 2C_{XX} + 2C_{XY} \leq N$ and with parameters $k_1 = 0.17$, $k_{-1} = 0.001$, $k_2 = 1.0$, $r = 1.001$. A maximal value $R_C = 0.00019$ was obtained. Even under the assumption of a carrying-capacity restriction (albeit physically questionable), the effect of stabilization (though rather weak for a site size of $N = 3$) in regions not attainable by the XY-system persists.

life. The prospects of supramolecular chemistry and there especially the usage of evolutionary mechanisms for adaptive molecular engineering [21] make an investigation of the details under which self sustained replication can take place for molecules which simultaneously carry mutable information and catalytic function of current technological interest.

This work was partially supported by the Swiss National Science Foundation and the German Research Council (DFG).

Appendix A

A central issue of the PRESS-framework is the limitation of the site size. As discussed, there is numerical evidence that even a moderate increase of N prohibits stabilization. This can be understood, when observing that there are no non-trivial stationary solutions for the CXY-system in a homogeneous environment, as is proven in the following.

In the homogeneous case, the dynamics of the CXY-system is given by a set of ODEs, which read (with $r_B = 1$, and the convention that small letter variables stand for the densities of the corresponding molecule type):

$$\begin{aligned}
\dot{x} &= 3k_2 c_{XX}\theta^2 + 2k_2(1-Q)c_{XX}\theta^2 + rk_2 c_{XY}\theta^2 \\
&\quad - 2k_1 x^2 - k_1 xy + 2k_{-1}c_{XX}\theta + k_{-1}c_{XY}\theta - dx, \\
\dot{y} &= k_2(1-Q)c_{XX}\theta^2 + 2rk_2 c_{XY}\theta^2 - k_1 xy \\
&\quad + k_{-1}c_{XY}\theta - dy, \\
\dot{c}_{XX} &= -k_2 c_{XX}\theta^2 + k_1 x^2 - k_{-1}c_{XX}\theta - dc_{XX}, \\
\dot{c}_{XY} &= -k_2 r c_{XY}\theta^2 + k_1 xy - k_{-1}c_{XY}\theta - dc_{XY}. \quad (9)
\end{aligned}$$

Here the sum of the densities is normalized to one, hence the density θ of the unoccupied places $\varnothing$ obeys $\theta = 1 - x - y - c_{XX} - c_{XY}$. In searching for stationary solutions, the time derivatives are set to zero, and the system can be analyzed further. The third and the forth equation are solved for c_{XX} and c_{XY}, respectively. The results are substituted in the first two equations, and they in turn are solved for the rates k_1 and k_{-1}. For k_{-1} one gets

see equation (10) above.

Taking into account that $Q < 1$ and $r \geq 1$ shows $k_{-1} < 0$, which is unphysical. Therefore, no stationary solutions exist for the complex forming, generous replicator model.

References

1. M.C. Bjoerlijst, P. Hogeweg, Physica D **48**, 17 (1991)
2. J.S. McCaskill, *Inaugural Lecture at Friederich Schiller University, Jena, Germany* (IMB Press, Jena, Germany, 1994)
3. J.S. McCaskill, R.M. Füchslin, S. Altmeyer, Biol. Chem. **382**, 1343 (2001)
4. M. Eigen, Naturwissenschaften **58**, 465 (1971)
5. M. Eigen, J.S. McCaskill, P. Schuster, Adv. Chem. Phys. **75**, 149 (1989)
6. P. Schuster, Biol. Chem. **382**, 1301 (2001)
7. D.S. Wilson, J.W. Szostak, Annu. Rev. Biochem. **68**, 611 (1999)
8. M. Eigen, P. Schuster, Naturwissenschaften **64**, 541 (1977)
9. J. Maynard Smith, Nature **280**, 445 (1979)
10. C. Bresch, U. Niesert, D. Harnasch, J. Theor. Biol. **85** 399 (1980)
11. M. Eigen, *Stufen zum Leben* (Piper, Munich, Germany, 1987)
12. B. Böddeker, diploma thesis, University of Göttingen, Germany (1995)
13. J.S. McCaskill, Biophys. Chem. **66**, 145 (1997)
14. S. Nee, Phil. Trans. R. Soc. Lond. B **355**, 1607 (2000)
15. R.M. Füchslin, J.S. McCaskill, Proc. Natl. Sci. Acad. **98**, 9185 (2001)
16. M. Kimura, Proc. Natl. Acad. Sci. **80**, 6317 (1983)
17. E. Szathmary, L. Demeter, J. Theor. Biol. **128**, 463 (1987)
18. E. Zintzaras, M. Santos, E. Szathmary, J. Theor. Biol. **217**, 167 (2002)
19. B. Drossel, Adv. Phys. **50**, 209 (2001)
20. E. Baake, W. Gabriel, Ann. Rev. Comp. Phys. **VII**, 209 (2000)
21. J.M. Lehn, Proc. Natl. Sci. Acad. **99**, 4763 (2002)