Escalation of error catastrophe for enzymatic self-replicators

BENEDIKT OBERMAYER and ERWIN FREY[(a)]

Arnold-Sommerfeld-Center for Theoretical Physics and Center for NanoScience, Ludwig-Maximilians-Universität München, Theresienstr. 37, 80333 München, Germany

Abstract. - It is a long-standing question in origin-of-life research whether the information content of replicating molecules can be maintained in the presence of replication errors. Extending standard quasispecies models of non-enzymatic replication, we analyze highly specific enzymatic self-replication mediated through an otherwise neutral recognition region, which leads to frequency-dependent replication rates. We find a significant reduction of the maximally tolerable error rate, because the replication rate of the fittest molecules decreases with the fraction of functional enzymes. Our analysis is extended to hypercyclic couplings as an example for catalytic networks.

Introduction. – According to the RNA world hypothesis [1], prebiotic biochemical life is thought to have merged through four steps: starting from the primordial on-enzymatic synthesis of nucleotides and their subseuent non-enzymatic polymerization into random RNA, hich in a third step would non-enzymatically replicate, atural selection would finally produce a set of funcional RNA enzymes (ribozymes), establishing exponenial growth and initiating RNA evolution. Despite considrable experimental progress [2, 3], as of today no truely self-replicating system has been evolved according to this hypothetic schedule. To assess its intrinsic plausibility, theory has mainly focused on the third step, usually based on the Eigen model [4] for prebiotic evolution: here, autocatalytic self-replication of L-nucleotide sequences proceeds non-enzymatically via stepwise template-directed polymerization, with a non-negligible error probability μ per single nucleotide. Assuming that one specific "master" template replicates with the highest rate $\alpha > 1$, while all other sequences have unit replication rate, it is found that faithful replication of the master is possible only for error probabilities smaller than a critical value $\mu_c \approx \ln \alpha / L$. In this regime, the population in sequence space is concentrated about the master in a rather broad distribution, giving rise to the notion of a "quasispecies". Larger values $\mu > \mu_c$ lead to a delocalized state with completely random sequences in the population. Many aspects of the

Eigen model depend to a large extent on the chosen fitness landscape, which assigns replication rates to genotypes. In the case of RNA, it displays a considerable degree of neutrality, because the mapping of sequences to secondary structures is decidedly many-to-one [5]. Still, although not universal, the existence of a critical mutation rate μ_c is a comparatively robust phenomenon [6, 7]. It has been termed "error catastrophe" [8], because it puts possibly irreconcilable simultaneous constraints on maximally tolerable error probability and minimal functional sequence length.

Lacking actual observations of freely self-replicating RNA and hence reliable estimates for replication rates, these theoretical limitations of non-enzymatic RNA replication are not yet reasonably quantitative. However, biochemical issues [9] raise severe doubts about its plausibility as well. Although ribozymes have been discovered that catalyze most of the necessary reaction steps [2,3,10,11], it remains questionable how a ribozyme should literally copy itself [10, 12]. Enzymatic replication seems the far more likely scenario, in the sense that a ribozyme copies other molecules. Presumably and most effectively, it would copy only those molecules that are exact replicas of itself, not only because known ribozymes act very substrate-specific, but also because unspecific recognition does not give a selective advantage to the replication enzymes themselves; this would require compartmentalization in vesicles to keep closely related molecules together [12]. Further, it
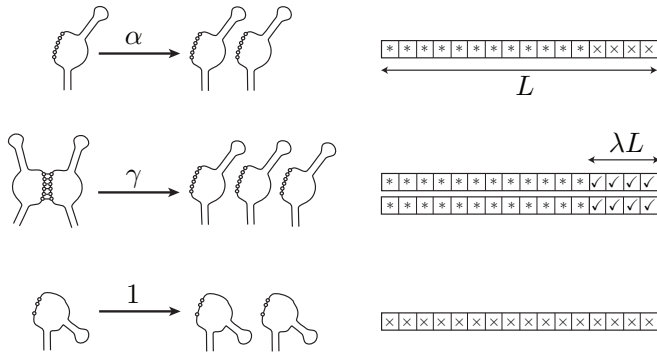
Fig. 1: Schematic illustration of the model. Left: Molecules with correct structure can replicate non-enzymatically with rate $\alpha > 1$. These molecules can also replicate enzymatically with rate γ, if they bind specifically to an identical partner within an otherwise selectively neutral recognition region of λL sites (dots). Misfolding mutant molecules replicate with unit rate. This model is formulated in terms of sequences rather than structures, as shown in the right panel: we distinguish correct "structural" nucleotides ($*$), matching sites in the recognition region ($\checkmark$) and unmatching or random nucleotides ($\times$).

has recently been suggested that the spontaneous emergence of RNA polymerases even without previous non-enzymatic replication could be promoted by a significant increase of functional complexity in a pool of random RNA due to the likely appearance of ligase activity [13].

In order to comparatively analyze non-enzymatic and enzymatic replication, their competition and their respective tolerance against mutations theoretically and by means of stochastic simulations, we employ a simplified quasispecies model, where sequences replicate both non-enzymatically and enzymatically, the latter with high specificity. We find a coexistence regime of these two replication modes, and an escalation of the error catastrophe in the enzymatic case: because the replication rate of the fittest molecules decreases with the fraction of functional enzymes, the maximally tolerable mutation rate is significantly reduced. To make contact to models of modular evolution and catalytic networks, where complex function is assumed to emerge through independent selection of small functional motives, thereby circumventing the error catastrophe [14, 15], we then extend our analysis to the case of hypercycles [8, 16, 17].

Model. – Motivated by the observation that catalytic and recognition regions are often clearly separated in ribozymes like the RNA component of RNaseP [11], we assume that the specific recognition mediating enzymatic replication involves only a small fraction λ of otherwise selectively neutral sites. This means that the majority of sites forms the proper secondary structure of the molecule and builds its active center, which catalyzes the polymerization reactions. Although secondary structure folding algorithms provide an improved genotype-fitness mapping

through an excellent approximation to RNA phenotypes, our model is formulated in terms of sequences instead of structures to allow for analytical treatment. We hence distinguish between "structural sites" and a "recognition region" on the sequence level (see Fig. 1 for a schematic illustration of our model). For the former, we use a sharply-peaked fitness landscape: a master sequence S^* has the highest non-enzymatic replication rate $\alpha > 1$, while all other sequences replicate with unit rate defining the time scale. We ignore possibly neutral sites in the structural region, because on our level of approximations this merely renormalizes their total number, or, equivalently, the mutation probability (see below). However, we do account for mutations in the recognition region, which do not affect non-enzymatic replication but the specificity of enzymatic replication: idealizing "highly specific", we require the recognition regions of enzyme and substrate to be identical for enzymatic replication to take place. Hence, ribozymes replicate only exact copies of themselves, with γ the associated rate constant. Note that we do not make any restrictions on the specific sequence of the recognition region: any molecule with the correct sequence for the structural sites can replicate enzymatically if it recognizes a suitable enzyme.

In the following, we formalize this model in the framework of quasispecies theory [4], where molecules are represented by sequences $S_i = (\sigma_1^{(i)} \sigma_2^{(i)} \ldots \sigma_L^{(i)})$ of L binary nucleotides $\sigma_j \in \{0, 1\}$. Their concentrations x_i evolve in the L-dimensional hypercube according to the deterministic rate equations

$$\dot{x}_i = \sum_k m_{ik} r_k x_k - x_i \bar{r}, \qquad (1)$$

where r_k is the replication rate of S_k, $m_{ik} = \mu^{d_{ik}}(1 - \mu)^{L-d_{ik}}$ is the mutation probability between sequences S_i and S_k with Hamming distance d_{ik}, and μ is the single-nucleotide mutation probability. The second term in Eq. (1) involves the mean replication rate $\bar{r} = \sum_k r_k x_k$ and ensures the normalization $\sum_k x_k = 1$. According to the above defined model, the replication rates read

$$r_k = \begin{cases} \alpha + \gamma x_k, & \text{if } S_k|_{\text{struc}} = S^*|_{\text{struc}}, \\ 1, & \text{otherwise.} \end{cases} \qquad (2)$$

In Eq. (2), $S_k|_{\text{struc}}$ denotes the restriction of the sequence S_k to the structural sites, and $S^*|_{\text{struc}}$ is the corresponding master sequence. While replication rates are usually taken as functions only of the genotype, with one single peak at the master sequence [4, 6, 7, 18–20], our model leads to frequency-dependent selection, which has only rarely been analyzed because it leads to mathematically challenging replicator-mutator equations (see, e.g., Ref. [21]).

Stochastic simulation. – For a realization of the full 2^L-dimensional system Eq. (1) in a finite population of N sequences, we employ the straightforward stochastic simulation algorithm used in Ref. [25]. At each time
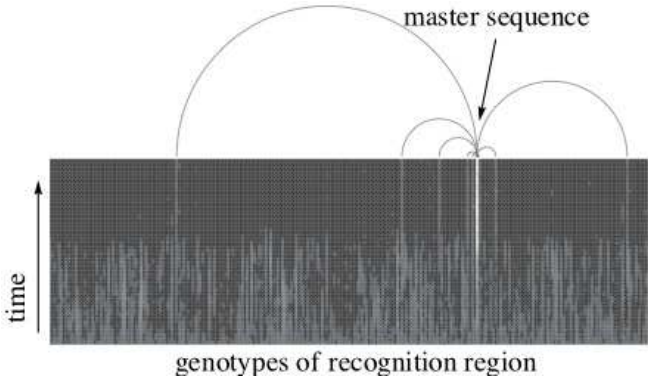
Fig. 2: Exemplary run of the stochastic simulation in a population of $N = 10^3$ sequences for $L = 32$, $\alpha = 5$, $\gamma = 10$, $\lambda = 1/4$, and $\mu = 0.005$. All sequences have been initialized with correct structural region but random recognition region. Their concentration is shown in gray level as function of time and genotype in the recognition region (linearly arranged by reading bit strings as integer numbers). Spontaneous concentration fluctuations lead to the establishment of a quasispecies of enzymatic replicators centered about one specific yet randomly chosen master sequence. Neighboring sequences are indicated by thin lines.

t each sequence S_k, present in n_k copies, has a probability $p_{0,k} = n_k / \sum_i n_i (1 + r_i)$ to be copied without mutations into the population at time $t + 1$, and a probability $p_{\mathrm{mut},jk} = m_{jk} r_k n_k / \sum_i n_i (1 + r_i)$ to be selected and mutated into sequence S_j. The population is initialized uniformly at the master sequence, but with random recognition sequences. Because the sites of the recognition sequence are effectively neutral, a state with all possible recognition sequences present in equal concentration is stable for large sequence length [21]. But if number fluctuations sufficiently increase the concentration of one particular *yet randomly chosen* sequence, this conveys via Eq. (2) a selective advantage, and its concentration will thus increase, up to the extent that mainly this sequence and its next mutational neighbors are present, in a quasispecies distribution very much like the one obtained in usual fitness landscapes. Fig. 2 shows an example of this outcome, which is somewhat reminiscent of a fixation event. While its detailed dependence on the specific formulation of the underlying stochastic process and the parameter values is left for future research, the localization itself turns out to be a robust phenomenon. In the following, we will therefore without loss of generality assume that the recognition region of the most populated sequence is equal to the one of the master sequence S^*.

Results. – While analytic solutions to the 2^L-dimensional system Eq. (1) are hard to obtain, we can use the so-called "error-tail" approximation [19]: here, we introduce three different classes of molecules. In x_{e}, we gather enzymatic replicators identical to the master sequence, with a replication rate $r_{\mathrm{e}} = \alpha + \gamma x_{\mathrm{e}}$. We use a second class x_{n} for non-enzymatic replicators, with struc-

tural sites identical to the master sequence but random recognition sequences. Their replication rate is $r_{\mathrm{n}} = \alpha$: although they are capable of enzymatic replication, the fraction of suitable enzymes with the appropriate recognition region is negligible. Finally, $1 - x_{\mathrm{e}} - x_{\mathrm{n}}$ is the error-tail of molecules with incorrect structural sites and unit replication rate. The main approximation of the error-tail approximation is to consider only those mutations that lead into a less-fitter class, with the probability not to have such a mutation abbreviated as "quality factor" Q. This approximation is generally valid for large sequence length but may fail if peaks in the fitness landscape are very dense [7]. The enzymatic replicators in x_{e} have $Q_{\mathrm{e}} = (1 - \mu)^L \equiv Q$, because a single error in L nucleotides suffices to destroy either structural or recognition region. The non-enzymatic replicators in x_{n} have a larger quality factor $Q_{\mathrm{n}} = (1 - \mu(1 - \lambda))^L \approx Q^{1-\lambda} > Q$: because the presence of λL neutral sites in the recognition region reduces the effective mutation probability, these sequences are mutationally more robust [22–24]. Further, with probability $Q^{1-\lambda} - Q$ mutations in x_{e} will hit a site of the recognition region and thus contribute to x_{n}. Hence, the dynamical system in the error-tail approximation is given by

$$\begin{aligned}
\dot{x}_{\mathrm{e}} &= r_{\mathrm{e}} Q x_{\mathrm{e}} - x_{\mathrm{e}} \bar{r} \\
\dot{x}_{\mathrm{n}} &= r_{\mathrm{n}} Q^{1-\lambda} x_{\mathrm{n}} + r_{\mathrm{e}} (Q^{1-\lambda} - Q) x_{\mathrm{e}} - x_{\mathrm{n}} \bar{r},
\end{aligned} \tag{3}$$

where the mean replication rate reads $\bar{r} = (r_{\mathrm{e}} - 1) x_{\mathrm{e}} + (r_{\mathrm{n}} - 1) x_{\mathrm{n}} + 1$. Solutions to the stationary state $\dot{x}_{\mathrm{e}} = \dot{x}_{\mathrm{n}} = 0$ of Eq. (3) for different mutation probabilities μ are shown in Fig. 3, together with results from a stochastic simulation of the full system with the replication rates Eq. (2) in a population of $N = 10^4$ sequences, where we initialized the sequences uniformly at the master sequence to reduce noise resulting from the intrinsically stochastic "fixation" events shown in Fig. 2, and averaged the results over time after reaching a stationary state. Obviously, approximating the deterministic rate equations with the simplified Eq. (3) gives an excellent description of the stochastic system. We can clearly distinguish three different regimes, separated by two error thresholds.

For high mutation probability, the population is delocalized over sequence space and only the error tail is significantly populated ($x_{\mathrm{e}} = x_{\mathrm{n}} = 0$). For smaller values of μ, we find a "non-enzymatic regime", where sequences with correct structural region are present, but a stable recognition sequence cannot be maintained, such that enzymatic replication is not possible. Explicitly, we find $x_{\mathrm{e}} = 0$ and $x_{\mathrm{n}} = (\alpha Q^{1-\lambda} - 1)/(\alpha - 1)$. The two regimes exchange stability at $Q = \alpha^{-1/(1-\lambda)}$, corresponding to $\mu = \mu_{\mathrm{c,n}} \approx \ln \alpha / (L(1 - \lambda))$. This is the familiar "phenotypic error threshold" [23,24]: the presence of neutral sites renormalizes the effective mutation probability, equivalent to having a shorter sequence [4].

For smaller mutation probabilities $\mu < \mu_{\mathrm{c,e}}$ the "enzymatic regime" becomes stable. Here, the fraction x_{e} of
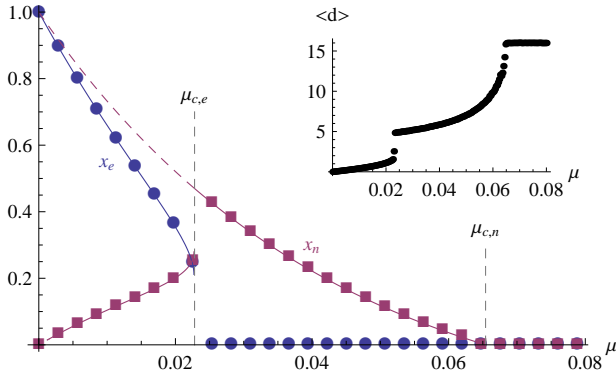
Fig. 3: Comparison between simulation results for x_e (circles) and x_n (squares) in a population of $N = 10^4$ sequences and solutions to Eq. (3) for $L = 32$, $\alpha = 5$, $\gamma = 10$ and $\lambda = 1/4$. The two error catastrophes occur at $\mu_{c,e} \approx \ln Q_c^{-1}/L$ with Q_c a solution of Eq. (4) and $\mu_{c,n} \approx \ln \alpha/(L(1-\lambda))$. The inset shows the average Hamming distance $\langle d \rangle$ to the master sequence, which increases in two steps, the first one at $\mu_{c,e}$ discontinuous, the second one at $\mu_{c,n}$ continuous.

enzymatic replicators is nonzero, but $x_n > 0$ as well, because this class is fed from x_e through mutations in the recognition region. Solving a third-order polynomial for x_e and x_n, we find this regime is stable when the corresponding discriminand is positive, which yields a critical value $Q = Q_c$ from the condition

$$4 \left[3\gamma(1 + \alpha Q_c(Q_c^{-\lambda} - 2)) - (\gamma Q_c + Q_c^{-\lambda} - \alpha)^2 \right]^3$$
$$+ \left[9\gamma(\alpha - \gamma Q_c - Q_c^{-\lambda})(1 + \alpha Q_c(Q_c^{-\lambda} - 2)) \right.$$
$$\left. + 27\alpha\gamma(\alpha Q_c - 1)(1 - Q_c^{-\lambda}) + 2(\gamma Q_c + Q_c^{-\lambda} - \alpha)^3 \right]^2 = 0. \tag{4}$$

Asymptotic solutions are given by:

$$Q_c \sim \begin{cases} 1 - \frac{\gamma}{4\alpha\lambda}, & \text{if } \gamma \ll \alpha, \\ \frac{2\sqrt{\gamma}-1}{\gamma} + \frac{\lambda}{\gamma} \ln \frac{2\sqrt{\gamma}-1}{\gamma} + \mathcal{O}(\lambda^2), & \text{if } \gamma \gg \alpha. \end{cases} \tag{5}$$

Note that the large-γ-limit is $Q_c \sim 2/\sqrt{\gamma}$, which implies for the corresponding critical value $\mu_{c,e} \approx \ln Q_c^{-1}/L \approx \ln \gamma/(2L)$. This significant reduction by a factor of 2 can be phrased as "escalation of error catastrophe": as the fraction x_e of enzymatically replicating sequences drops with higher mutation probability, their replication rate $r_e = \alpha + \gamma x_e$ decreases as well, leading to an even stronger reduction in x_e. Beyond the critical value $\mu_{c,e}$, the fraction of molecules with the correct recognition sequence becomes so small that their replication rate is not large enough for them to be maintained in the population at a macroscopic level.

An important difference between the transitions at $\mu_{c,n}$ and $\mu_{c,e}$ can be observed not only in the fraction x_e, but also in the width of the population distribution (measured
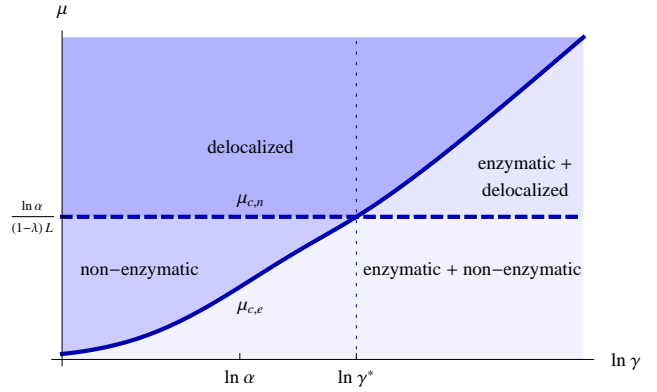

Fig. 4: Phase diagram of stability regimes of Eq. (3) in the $\ln\gamma$-μ-plane: the critical value $\mu_{c,n}$ (thick dashed line) separates the delocalized regime (above) and the non-enzymatic regime (below). Enzymatic replication is stable below $\mu_{c,e}$ (thick solid line) and becomes mutationally more robust than non-enzymatic replication if $\gamma > \gamma^* = \mathcal{O}(\alpha^2)$ (vertical line).

as average Hamming distance to the master sequence), shown in the inset of Fig. 3: while the delocalization transition at $\mu = \mu_{c,n}$ is continuous, the transition at $\mu_{c,e}$ is discontinuous. In the former case, this property depends also on the choice of observable [7], but in the latter case, the discontinuity results from bistability: together with the enzymatic regime, also the non-enzymatic regime or the delocalized phase may be stable, depending on whether μ is larger or smaller than $\mu_{c,n}$, and if $\mu > \mu_{c,e}$ the enzymatic regime vanishes. The phase diagram in Fig. 4 summarizes these various regimes. Note that $\mu_{c,e} = \mu_{c,n}$ at a critical value

$$\gamma^* = \alpha(\alpha - 1) + 2\lambda^{1/2}\alpha\sqrt{2\alpha(\alpha - 1)\ln\alpha} + \mathcal{O}(\lambda). \tag{6}$$

This result implies that very large rates $\gamma = \mathcal{O}(\alpha^2)$ are required if enzymatic replication is to be more error-tolerant than non-enzymatic replication [16]. Although this possibility is not contained in the approximate Eq. (3), we find that in the bistability region $\mu < \min(\mu_{c,e}, \mu_{c,n})$ the enzymatic regime is easily populated by selectively advantageous concentration fluctuations from the non-enzymatic regime by randomly choosing a "master" sequence for the recognition region as shown in Fig. 2.

Extension to hypercyclic couplings. – The realization that replication errors limit the maximum complexity of self-replicating molecules to a possibly paradoxical extent has lead to theories of *modular* evolution, where complex functions emerge through catalytic interactions of smaller independently selected motifs [14,15], thereby also speeding up evolution by facilitating the search for complexity. While arbitrarily complex interaction networks between different modules or molecular species are conceivable, the simplest case applicable to the above system with its two-molecule interactions is the hypercycle [8]. Here, n species are arranged in a circular directed graph,

where each species enzymatically catalyzes the replication of its next neighbor. This network gives rise to coexistence of all species, in a stable fixed point for $n \leq 4$ and via periodic orbits for larger n. In contrast to previous approaches accounting for replication errors in a hypercycle [16,17], we consider distinct error tails for all species: each is present in an enzymatically active variant $x_{e,i}$ with replication rate $r_{e,i} = \alpha_i + \gamma_i x_{e,i+1}$ together with its non-enzymatic error tail $x_{n,i}$ with replication rate $r_{n,i} = \alpha_i$. In addition, there is the global error tail of misfolding mutants. For simplicity, we assume a symmetric setup with identical rate constants $\alpha_i \equiv \alpha$ and $\gamma_i \equiv \gamma$. This gives the rate equations

$$
\begin{aligned}
\dot{x}_{e,i} &= r_{e,i}Q x_{e,i} - x_{e,i}\bar{r} \\
\dot{x}_{n,i} &= r_{n,i}Q^{1-\lambda} x_{n,i} + r_{e,i}(Q^{1-\lambda} - Q)x_{e,i} - x_{n,i}\bar{r},
\end{aligned}
\tag{7}
$$

where indices are taken modulo n and the mean fitness is now given by $\bar{r} = (\alpha-1)\sum_i (x_{e,i}+x_{n,i})+\gamma\sum_i x_{e,i}x_{e,i+1}+1$. It is easy to see that this system reduces to Eq. (3) if we assume that $x_{e,i} \equiv x_e^*/n$ and $x_{n,i} \equiv x_n^*/n$ and replace $\gamma \to \gamma/n$. Then, the "enzymatic" solution of Eq. (3) corresponds to the inner fixed point of the hypercycle, where all species are present in equal concentration. Replication errors can be tolerated only if $\mu < \mu_{c,e}$, where $\mu_{c,e} \sim \ln(\gamma/n)/(2L)$ for large γ (see also Ref. [16]). We emphasize that in contrast to Ref. [14], where unspecific ligation was assumed to link functional motifs, in our model specific recognition between different species leads to frequency-dependent replication rates. This reduces the error threshold by roughly a factor of 2, which in a two-member hypercycle would cancel the putative complexity gain resulting from using two subunits of half the sequence length.

Moreover, increasing the number of hypercycle members beyond $n = 4$ changes the stability of the central fixed point. Observing that the Jacobian matrix of Eq. (7) is block-circulant [26] (every block is a 2×2-matrix for the two concentration variables $x_{e,i}$ and $x_{n,i}$ per species), its crucial eigenvalues with possibly non-negative real part are given by $\frac{1}{n}\gamma Q x_e^* e^{im\pi/n}$, where $m = 0,\ldots,n-1$ and x_e^* is the enzymatic solution of Eq. (3) with $\gamma \to \gamma/n$. Hence, these eigenvalues are proportional to the n different nth roots of unity. In close correspondence to the error-free hypercycle (and in contrast to Ref. [16], where a stability region for $n = 5$ was found), the central fixed point loses stability for $n > 4$, giving rise to limit cycles with large concentration oscillations, which are vulnerable to extinction via stochastic fluctuations.

Note that the stable inner fixed point corresponding to the enzymatic regime implies coexistence of different species. However, in the non-enzymatic regime of Eq. (3), the different error tails do compete against each other. As soon as the hypercycle breaks down, e.g., because the recognition sequence is lost due to stochastic fluctuations, one error tail will drive the others to extinction, due to the competitive exclusion principle encountered in usual

quasispecies theory [4]. This makes the reverse process, i.e., a fluctuation that establishes a closed cycle, extremly unlikely.

Conclusion. – In summary, we have analyzed a simple quasispecies model for the non-enzymatic and enzymatic replication of ribozymes, where specific recognition is mediated via otherwise neutral sites. We find that the frequency-dependent replication rates associated with specific enzymatic replication lead to a discontinuous transition at the error threshold due to bistability with a partly delocalized phase. Further, hypercyclic couplings enable coexistence of at most four different species and their respective error tails in a stable fixed point.

$$* * *$$

REFERENCES

[1] ORGEL L., *Crit Rev Biochem Mol* , **39** (2004) 99.
[2] JOHNSTON W., UNRAU P., LAWRENCE M., GLASNER M. and BARTEL D., *Science* , **292** (2001) 1319.
[3] LINCOLN T. and JOYCE G., *Science* , **323** (2009) 1229.
[4] EIGEN M., McCASKILL J. and SCHUSTER P., *Adv. Chem. Phys.* , **75** (1989) 149.
[5] HUYNEN M., STADLER P. and FONTANA W., *Proc Natl Acad Sci USA* , **93** (1996) 397.
[6] WIEHE T., *Genet Res Camb* , **69** (1997) 127.
[7] JAIN K. and KRUG J., *Adaptation in simple and complex fitness landscapes* arXiv:q-bio/0508008 (2005).
[8] EIGEN M. and SCHUSTER P., *Naturwissenschaften* , **65** (1978) 7.
[9] ORGEL L., *Nature* , **358** (1992) 203.
[10] JOYCE G. F., *Angew. Chem.* , **46** (2007) 6420.
[11] LILLEY D. M., *Curr Opin Struc Biol* , **15** (2005) 313.
[12] SZOSTAK J., BARTEL D. and LUISI P. L., *Nature* , **409** (2001) 387.
[13] BRIONES C., STICH M. and MANRUBIA S., *RNA* , **15** (2009) 743.
[14] MANRUBIA S. C. and BRIONES C., *RNA* , **13** (2007) 97.
[15] TAKEUCHI N. and HOGEWEG P., *Biol Direct* , **3** (2008) 11.
[16] CAMPOS P., FONTANARI J. and STADLER P., *Phys. Rev. E* , **61** (2000) 2996.
[17] SILVESTRE D. A. M. M. and FONTANARI J. F., *Journal of Theoretical Biology* , **254** (2008) 804.
[18] SCHUSTER P. and SWETINA J., *Bull Math Biol* , **50** (1988) 635.
[19] NOWAK M. and SCHUSTER P., *J Theor. Biol* , **137** (1989) 375.
[20] SAAKIAN D. and HU C., *P Natl Acad Sci USA* , **103** (2006) 4935.
[21] STADLER P., SCHNABL W., FORST C. and SCHUSTER P., *Bull Math Biol* , **57** (1995) 21.
[22] WILKE C., WANG J., OFRIA C., LENSKI R. and ADAMI C., *Nature* , **412** (2001) 331.

[23] TAKEUCHI N., POORTHUIS P. and HOGEWEG P., *BMC Evol Biol* , **5** (2005) 9.

[24] KUN A., SANTOS M. and SZATHMARY E., *Nat Genet* , **37** (2005) 1008.

[25] WILKE C., RONNEWINKEL C. and MARTINETZ T., *Phys Rep* , **349** (2001) 395.

[26] TEE G., *Res. Lett. Math. Sci* , **8** (2005) 123.