Graded Artificial Chemistry in Restricted Boundaries

Barak Shenhav, Ran Kafri and Doron Lancet

Dept.of Molecular Genetics and the Crown Human Genome Center,
The Weizmann Institute of Science, Rehovot 76100, Israel
{barak.shenhav, ron.kafri, doron.lancet}@weizmann.ac.il

Abstract

The question of the origin of life is addressed by artificial life research, particularly in the realm of artificial chemistry. Such artificial chemistry is described by our Graded Autocatalysis Replication Domain (GARD) model. GARD depicts an unorthodox scenario suggested for emergence of life – the 'lipid world'. The model concerns molecular assemblies with mutual catalysis in an environment containing a plethora of molecular species. Many aspects of GARD were amply discussed. Here we concentrate on the importance of size constraints as depicted by the basic model and several of its variants. Occasional fission of a GARD assembly, which restricts the assembly size, is crucial for generating compositional quasi-stationary states ('composomes'). In a spatial version of GARD, bounded environments yield spontaneous emergence of different ecologies. Limiting the size of a population of GARD assemblies gives rise to a complex population dynamics. The last example, with possible wider impact to chemistry and nano-technology, suggests that size limit can give rise to spontaneous symmetry breaking. This latter result is compared to the classic Frank's model for homo-chirality, which requires explicit inhibition. We conclude that size restrictions are fundamental in the field of origin of life and artificial life, not only in order to facilitate evolutionary processes, as previously suggested, but also, for augmenting the dynamics portrayed by different scenarios and models.

Introduction

The question "How does life arise from the nonliving?" is first on the list of open problems in artificial life [1]. The study of the origin of life is traditionally considered to be related to classic chemistry. Starting from the seminal experiments of Oparin and Miller, a large body of knowledge was gathered on the synthesis of organic matter under possible prebiotic conditions. Combined with later insight regarding the likely supply of compounds from extraterrestrial sources, it seems that the question of availability of basic building blocks for early life is largely resolved, though many of the details are still missing, such as the exact nature of the molecules, their quantities or the environmental parameters (e.g. temperature, pressure, pH).

The assembly or emergence of life, given that the right building blocks do exist, is still out of the scope of experimental chemistry. The question of organization in a seemingly random chemical scenario was tackled by theoretical models, in the realm of artificial chemistry [2], such as the Quasi-species and the hyper-cycle models [3, 4]. Strongly related to the work presented hereafter are the models depicting the *metabolism first* [5] approach concerning the emergence of autocatalytic sets [6, 7] and their capacity to evolve [8, 9]. These models use very simplified artificial chemical rules, usually with binary (all or nothing) parameters.

The Graded Autocatalysis Replication Domain (GARD) model takes an intermediate approach. While due to the unavailability of details, GARD is comprised of abstract molecules, it employs rigorous chemical kinetics equations and employs parameters from a realistic chemical probability distribution.

The Model

GARD entails non-covalent assemblies of mutually catalytic molecules. The molecules are usually thought to be amphiphiles (lipids), forming spontaneous molecular assemblies, e.g. micelles or liposomes, as suggested by the 'lipid world' scenario for origin of life [10]. Under the choppy conditions on early earth, such assemblies are expected to undergo occasional fission, which can serve as a primordial progeny-generation mechanism. This process presents the transfer of *compositional information* from a parent assembly to daughter assemblies as discussed in details in previous work [11, 12].

Model Description

The basic GARD model describes a single assembly in an environment containing a finite molecular repertoire of size N_G. The external concentration of the molecules, ρ, is generally taken to be equal for all species. The internal molecular counts are n_1, n_2 …, (or in vector form $\mathbf{n} = [n_1 \quad ... \quad n_{N_G}]$). Thus, the assembly size is $N = \sum n_i$. Molecules join and leave the assembly with spontaneous rates k_f and k_b, typically taken to be equal for all molecular species. The crux of the model lies in the introduction of catalytic rate enhancements, exerted by molecules within the assembly on the join/leave reactions. Using a statistical chemistry approach, the values of the rate enhancement exerted by molecules of type j on molecules of type i (β_{ij}) are drawn from a lognormal distribution. This distribution is based of the Receptor Affinity Distribution (RAD) [13] modified for rate enhancement [12, 14].

The characteristic equations that describe an assembly's growth, in terms resembling continuous concentrations, are [11]:

$$\frac{dn_i}{dt} = \left(k_f \rho N - k_b n_i\right)\left(1 + \frac{1}{N}\sum_{j=1}^{N_G}\beta_{ij} n_j\right) \qquad (1)$$

For most computer simulations of GARD, discrete stochastic chemistry algorithms were used. Stochastic chemistry is suitable for small molecular systems because it reflects stochastic dynamic noise and granularity as opposed to the common differential equation approach. Early works used the 'τ-leap' algorithm [15]. Later papers [16, 17] employed the older, yet (for the special case of GARD) much faster, 'first reaction method' [18, 19]. These calculated the reaction rates by:

$$J_i = k_f e_i\left(1 + \sum_{j=1}^{N_G}\beta_{ij} n_j\right) \qquad (2a)$$

$$L_i = k_b n_i\left(1 + \sum_{j=1}^{N_G}\beta_{ij} n_j - \beta_{ii}\right) \qquad (2b)$$

where J_i and L_i are, respectively, the join and leave rates of molecular species i, and e_i is its external count near the assembly. The subtraction of β_{ii} in eq. 2b depicts that a molecule cannot exert catalysis on itself, an idea generally ignored when concentrations are considered.

Under conditions of unlimited external resources, the molar fractions n_i/N reach a single stationary state **n***, which corresponds to the eigen-vector with the highest eigen-value of a linear form of eq. 1 [11, 20]. Moreover, if the external supply of molecules is limited, the system will reach equilibrium for which n_i corresponds to the ratio of k_f and k_b. However, when the growth is interrupted by occasional fission, (or split), the dynamics observed may be altered dramatically. A split is modeled by random removal of molecules from the assembly, once it reaches a critical size N_{MAX}.

The Major Result - Composomes

In a system consisting of a single assembly, when a split is introduced, the effect of limited external resources is generally not observed, as the environment is generally not depleted. Thus, a GARD system is expected to approach the canonic stationary state **n***. Indeed, for the case where the size of the assembly at the split is large ($N_{MAX} >> N_G$) the assembly assumes a composition near the canonic one (figure 1a).

When $N_{MAX} \approx N_G$, often several quasi-stationary states – *composomes* are observed (figure 1b). Typically, one of the composomes corresponds to **n***, though this does not have to be the most frequent one. Each composome, in this case, arises from a different underlying catalytic network [11]. These networks show mutual catalytic closure [6] to some extent, which gives them profound stability against destructive splits.

For the extreme case where $N_{MAX} << N_G$ often the assembly drifts most of the time in a non-stationary state. Yet, numerous short lasting quasi-stationary states are perceived, corresponding to the temporary existence of a relative strong catalyst in the assembly.

To summarize, GARD assemblies do not converge to a single steady state when a size limit is imposed on them by random fission.

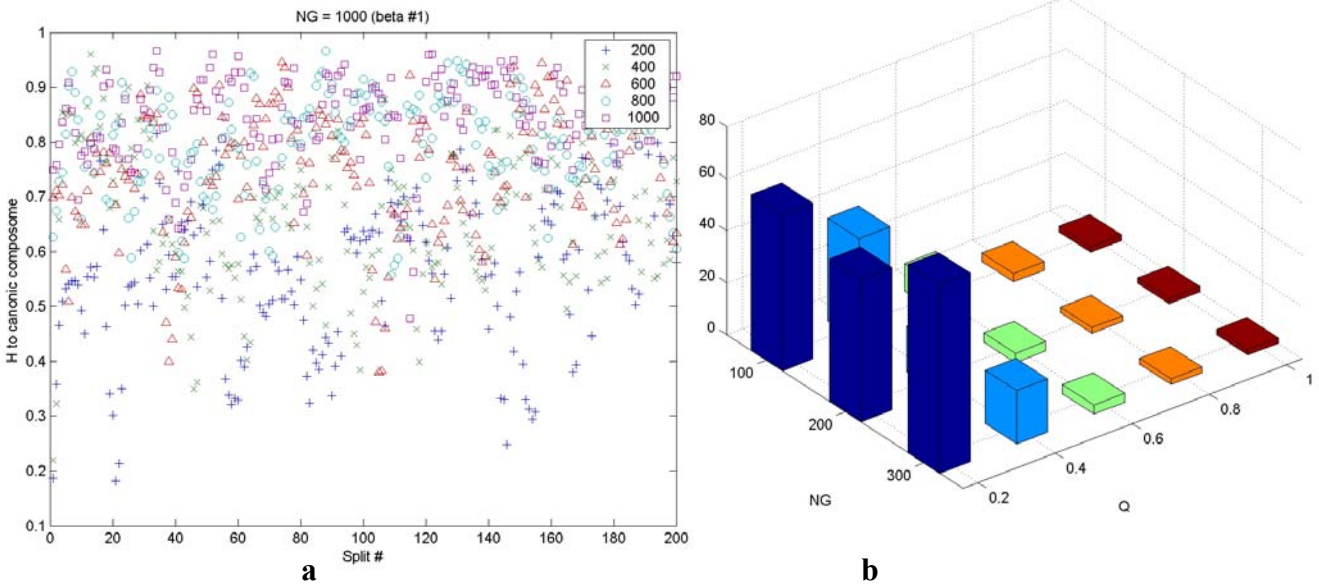

Figure 1 – Behavior of basic GARD assemblies. (a) Similarity to the canonic composition **n*** (as defined in the text) in 3 different runs of GARD for 200 splits each. All runs use the same values for k_f, k_b,β and ρ. N_G=1000 and N_{MAX} differ (200, 600, 1000). (b) The average number of composomes observed in runs of GARD for 2000 splits each with different N_G and different Q (= N_G/N_{MAX}) values. Each bin was computed by averaging 30 runs. Composomes were found using the 'on-the-fly' algorithm described in [17].

GARD Extensions

Recently, we have begun to explore more realistic, and generally more computationally demanding, extensions of the GARD models.

Polymer GARD (P-GARD) introduces additional reactions, which allow the formation of more complex molecules. These reactions mimic the formation of chemical covalent bonds between two molecules within the assembly. In a first study of the model [17], with solely monomers (basic molecular species) and dimers (a concatenation of two monomers), we have observed composomes that have appreciable dimer content, and the appearance of novel metabolism-like networks for internal dimer synthesis.

The Computational Origin of Life Endeavor (CORE) [21] aims, by means of extensive computational power, to turn GARD simulations as realistic as feasible. This includes, as a first step, large molecular repertoires and P-GARD simulations. A further step would be to use exact molecular dynamics simulations, rather than the statistical chemistry approach, for obtaining parameters values. This will allow GARD simulations with concrete molecules in place of the abstract ones used contemporary.

Chiral GARD (C-GARD) (Ron Kafri *et al.* – in preparation) imposes restrictions on the values of β_{ij} to reflect that most molecular species have two chiral enantiomers (D, L). According to 'Wigner's rule', D-enatiomers of a molecular species i interact with D-enatiomers of molecular species j in the same manner that L-enatiomers of molecular species i interact with L-enatiomers of molecular species j, that is:

$$\beta_{i_D j_D} = \beta_{i_L j_L} \qquad (3a)$$

The rule also postulates that:

$$\beta_{i_L j_D} = \beta_{i_D j_L} \qquad (3b)$$

With such β values, when composomes are observed, they tend to be either symmetric (racemic), containing equal quantity of both ennatiomers for each molecular species, that is:

$$n_{i_L} \approx n_{i_D} \quad i = 1, ..., N_G \qquad (4a)$$

or asymmetric, e.g.

$$n_{i_L} \approx 0 \, , \, n_{i_D} > 0 \qquad (4b)$$

GARD Populations explores the spatio-temporal development of populations of assemblies. This is investigated in a direct manner with a computer simulation consisting of many assemblies scattered in a spatial grid [16] and with more abstract formalism [17] (further detailed in Arren Bar-Even *et al.* – in preparation). This

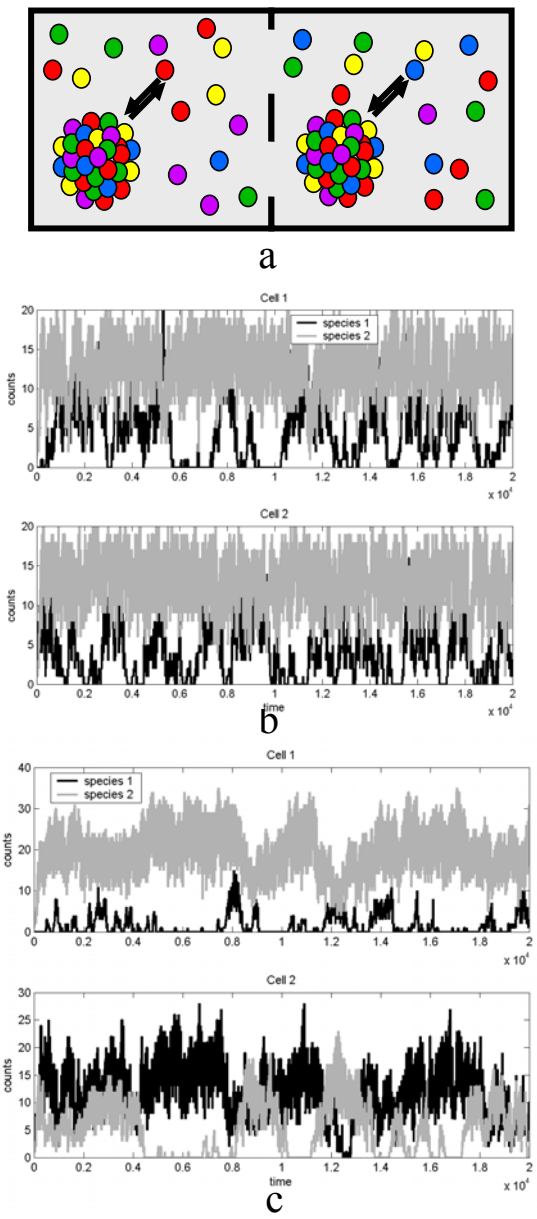

Figure 2 – The two assemblies two environment model. (a) A schematic description of the model consisting of N_G different molecular species (different shaded circle), two assemblies each dwelling in its own environment (different squares). Molecules in each environment can join/leave the assembly (depicted by the arrows). (b) The dynamics of the two assemblies with no diffusion between the environments. Light gray indicates the number of molecules supporting composome 1 (the dominant one) and Black the number of molecules supporting composome 2 (the second best composome). At the beginning, there are 20 copies of each molecular species at each environment . (c) same as case b with diffusion.

quest sheds light on emergence of rudimentary ecologies of mesobiotic entities, intermediates on the path between

inanimate prebiotic chemical entities to full-fledged (biotic) living cells [22].

Results

While the basic GARD model governs the framework of our efforts, we have considered several simple variants of it, which provide insights on different aspects of the behaviors observed.

Two Environments – Two Assemblies

An example for such GARD variant is a simple population model consisting of only two assemblies, each dwelling in its own external environment (figure 2a). The rules governing the dynamics of the assemblies are similar to the rules for a single assembly described above (eqs. 2a, 2b). If the environments are independent, then the dynamics of each assembly resembles that of a single assembly.

A more interesting case is when the two environments do interact, for example, by allowing diffusion between them. In many cases, such as when the environments consist of large amounts of each molecular species or when the diffusion is fast, both environments tend to portray similar dynamics (figure 2b). Yet, if the environment size is small compared to the size of the assemblies, an initial symmetric distribution of molecules between the environments may be spontaneously broken. To be exact, one of the assemblies assumes a certain composome and molecules constituting this composome concentrate in one compartment, while the other assembly assumes another composome, leading to the concentration of other molecular species in the other compartment (figure 2c).

This result are akin, to some extent, to the results reported for many assemblies in a single environment (Lipidia) [16] showing that limited resources lead to deviation from typical dynamics. Yet, there are differences in mechanism and in outcome: spatial organization rather than speciation in Lipidia. The phenomena observed amount to induced environmental inhomogeneity, rather than merely a stochastic one.

Population Catastrophes

The formalism developed to investigate populations of GARD assemblies treats composomes as abstract entities with emergent properties such as the time elapsing between assembly divisions, the probability of an assembly to preserve its composomal state following a split or the emergence likelihood of the composome [17, 23]. The use of such abstract entities, rather than detailed molecular assemblies, has enabled easy computer simulations of populations of assemblies. In the case where the population size was not bounded, a single stationary population distribution was observed, in similarity to the canonic stationary state **n*** described above. In contrast, several quasi-stationary population distributions were observed in cases where the constant population size was imposed [24] or where populations went through occasional 'catastrophes' ('killing' half of the population once it reached a threshold size).

Detailed analysis of this result and other insights coming up from this model are out of the scope of the current manuscript and are described elsewhere (Arren Bar-Even *et al.* – in preparation). Yet, it is another example that restricting the dynamics to a limited size alters the system behavior.

Chiral toy model

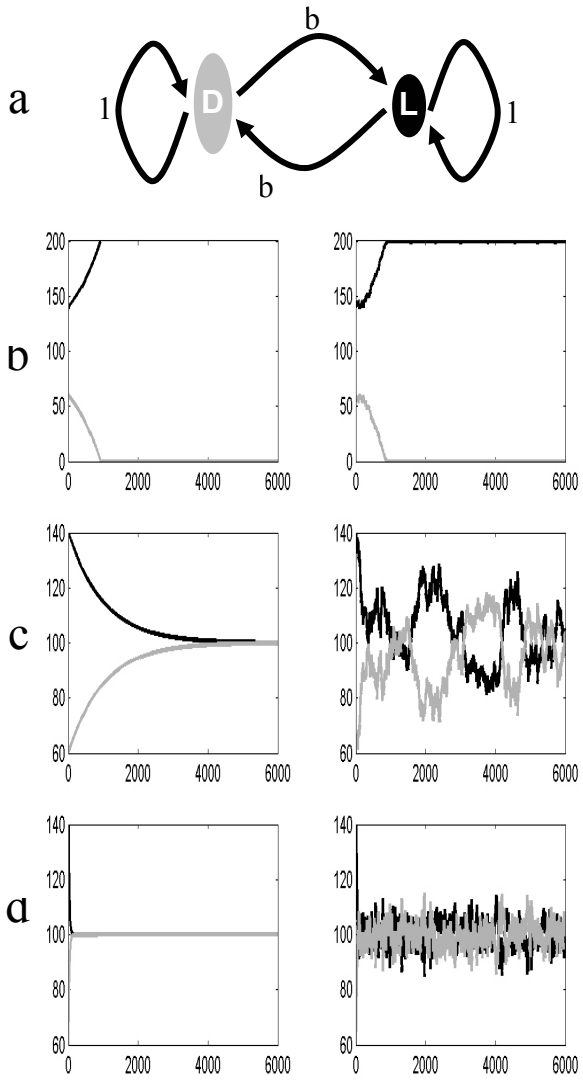

Figure 3 – The chiral toy model. (a) A schematic description of the model consisting of an L and D molecules with auto-catalysis of unit size and cross-catalysis of b. (b-d) The fraction of L molecules (Dark) and D molecules (Light) as function of time with different b values. Left shows the continuous case with no size limit (Frank's model) and right shows the stochastic dynamics with split (the chiral toy model). The values of b are: -0.05, 0.05 and 2 respectively.

We have explored the simplest possible chiral GARD, one with two molecular species, a single L molecular species and its corresponding D molecular species. For this special case, there are two independent rate enhancement values – autocatalysis (β_{LL}) and cross-catalysis (β_{LD}) since from eq. (3a) $\beta_{DD} = \beta_{LL}$ and from eq. (3b) $\beta_{DL} = \beta_{LD}$. Selecting the right set of units, we can take the autocatalysis to be 1, without losing any generality, denoting the cross-catalysis with b (figure 3a). The growth equations of this abstract GARD model are thus reduced to:

$$\frac{dn_L}{dt} = \left(k_f e_L - k_b n_L\right)\left(1 + n_L + b n_D\right) \qquad (5a)$$

$$\frac{dn_D}{dt} = \left(k_f e_D - k_b n_D\right)\left(1 + n_D + b n_L\right) \qquad (5b)$$

Eqs. (5a, 5b) resemble those described in Frank's model for the spontaneous emergence of homo-chiral systems [25], whichcontain solely L molecules or solely D molecules. In order to observe such spontaneous symmetry breaking, Frank's model requires the cross-catalysis to be negative, i.e. inhibitory. Indeed, for the cases where $b < 0$ or $b > 1$ the current model and Frank's model show the same behavior up to stochastic noise (figure 3b, 3d). However, for $0 < b < 1$ the two models display different behaviors. In Frank's model a racemic system emerges, containing similar amounts of L and D molecules. In our toy chiral GARD model, which includes a size limit (N_{MAX}), temporary biases towards L or D (enantio excesses) are observed (figure 3c).

This model may serve as a basis for studies of enantio – selection in larger molecular repertoires. For example, an asymmetric multi-component composome may be analogous to a bimolecular case with dominant autocatalysis (b < 1), while a symmetric multi-component composome could correspond to a bi-molecular system in which cross-catalysis is stronger ($b > 1$). This appears to be valid only under limited size constraints.

Whether the model relates to a single D and L molecular species or to a multi-molecular composome, it highlights the impact of imposing size boundaries on the system, leading to temporal symmetry breaking without requiring explicit inhibition.

Discussion

The basic GARD model sheds light on many aspects usually not concerned in traditional origin of life studies. We have concentrated here on the impact of imposing size restriction on the system. We have indicated that in the basic GARD model, as well as in several of its derivatives, such restrictions significantly alter the ensuing dynamics. Generally, the dynamics observed is more multifaceted, e.g. showing more quasi-stationary states or spontaneous symmetry breaking. Thus, we suggest size limiting as a general mechanism for turning seemingly lifeless systems, ones that portray a random walk or convergence to a single steady state, into more elaborate systems with several or many quasi-stationary states, capable of manifesting more lifelike faculties.

Other studies have highlighted the importance of restricted boundaries for the facilitation of evolutionary processes. For example, compartmentalization was suggested in order to maintain hyper-cycles against possible molecular parasites [26]. Another example is the requirement of encapsulating molecular replication mechanisms (replicases) by lipid vesicles. This was suggested to serve a crucial rule in allowing the evolution and takeover of better replicases [27]. We propose that augmenting the dynamics of the systems is another important aspect of closure, which requires further study.

Acknowledgments

We thank Barak Naveh for critically reading the manuscript. The research was partially supported by the European Commission through grant number HPRI-CT-1999-00026 (the TRACS Programme at EPCC) and by the Crown Human Genome Center.

References

[1] Bedau, M.A., McCaskill, J.S., Packard, N.H., Rasmussen, S., Adami, C., Green, D.G., Ikegami, T., Kaneko, K. and Ray, T.S., (2000). Open Problems in Artificial Life. Artif Life 6:363-376.

[2] Dittrich, P., Ziegler, J. and Banzhaf, W., (2001). Artificial Chemistries - a Review. Artif Life 7:225 –275.

[3] Eigen, M., (1971). Selforganization of Matter and the Evolution of Biological Macromolecules. Naturwissenschaften 58:465-523.

[4] Eigen, M. and Schuster, P., (1979). The Hypercycle. Berlin: Springer Verlag.

[5] Morowitz, H.J., (2002). The Emergence of Metabolism. In The Emergence of Everything: How the World Became Complex. Oxford Press.

[6] Farmer, J.D., Kauffman, S.A. and Packard, N.H., (1986). Autocatalytic Replication of Polymers. Physica 22D:50-67.

[7] Bagley, R.J. and Farmer, J.D., (1991). Spontaneous Emergence of a Metabolism. In Langton, C.G., Taylor, C., Farmer, J.D. and Rasmussen, S., Artificial Life II. Addison-Wesley.

[8] Jain, S. and Krishna, S., (2002). Large Extinctions in an Evolutionary Model: The Role of Innovation and Keystone Species. Proc. Natl. Acad. Sci. U. S. A. 99:2055-2060.

[9] Bagley, R.J., Farmer, J.D. and Fontana, W., (1991). Evolution of a Metabolism. In Langton, C.G., Taylor, C., Farmer, J.D. and Rasmussen, S., Artificial Life II. Addison-Wesley.

[10] Segre, D., Ben-Eli, D., Deamer, D.W. and Lancet, D., (2001). The Lipid World. Origins Life Evol B 31:119-145.

[11] Segre, D., Ben-Eli, D. and Lancet, D., (2000). Compositional Genomes: Prebiotic Information Transfer in Mutually Catalytic Noncovalent Assemblies. Proc Natl Acad Sci U S A 97:4112-7.

[12] Segre, D., Shenhav, B., Kafri, R. and Lancet, D., (2001). The Molecular Roots of Compositional Inheritance. J Theor Biol 213:481-491.

[13] Lancet, D., Sadovsky, E. and Seidemann, E., (1993). Probability Model for Molecular Recognition in Biological Receptor Repertoires - Significance to the Olfactory System. Proc. Natl. Acad. Sci. U. S. A. 90:3715-3719.

[14] Lancet, D., Kedem, O. and Pilpel, Y., (1994). Emergence of Order in Small Autocatalytic Sets Maintained Far from Equilibrium - Application of a Probabilistic Receptor Affinity Distribution (Rad) Model. Ber. Bunsen-Ges. Phys. Chem. Chem. Phys. 98:1166-1169.

[15] Gillespie, D.T., (2001). Approximate Accelerated Stochastic Simulation of Chemically Reacting Systems. J. Chem. Phys. 115:1716-1733.

[16] Naveh, B., Sipper, M., Lancet, D. and Shenhav, B., (submitted). Lipidia: An Artificial Chemistry of Self-Replicating Assemblies of Lipid-Like Molecules. In Proceedings of Artificial Life IX, Boston, MA, USA. MIT press

[17] Shenhav, B., Bar-Even, A., Kafri, R. and Lancet, D., (in press). Polymer Gard: Computer Simulation of Covalent Bond Formation in Reproducing Molecular Assemblies. Origins Life Evol B

[18] Gillespie, D.T., (1976). General Method for Numerically Simulating Stochastic Time Evolution of Coupled Chemical-Reactions. J. Comput. Phys. 22:403-434.

[19] Gillespie, D.T., (1977). Master Equations for Random Walks with Arbitrary Pausing Time Distributions. Physics Letters A 64:22-24.

[20] Jain, S. and Krishna, S., (1998). Autocatalytic Sets and the Growth of Complexity in an Evolutionary Model. Phys. Rev. Lett. 81:5684-5687.

[21] Shenhav, B. and Lancet, D., (2004). Prospects of a Computational Origin of Life Endeavor. Origins Life Evol B 34:181-194.

[22] Shenhav, B., Segre, D. and Lancet, D., (2003). Mesobiotic Emergence: Molecular and Ensemble Complexity in Early Evolution. Adv Complex Syst 6:15-35.

[23] Bar-Even, A., Shenhav, B., Kafri, R. and Lancet, D., (in press). The Lipid World: From Catalytic and Informational Headgroups to Micelle Replication and Evolution without Nucleic Acids. In Seckbach, J., Cellular Origin, Life in Extreme Habitats and Astrobiology (Cole).

[24] Kuppers, B.O., (1983). Molecular Theory of Evolution. Berlin-Heidelberg: Springer-Verlag.

[25] Frank, F.C., (1953). On Spontaneous Assymetric Synthesis. Biochem Biophys Acta 11:459-463

[26] Eigen, M., Gardiner, W.C. and Schuster, P., (1980). Hypercycles and Compartments. J Theor Biol 85:407-411.

[27] Szostak, J.W., Bartel, D.P. and Luisi, P.L., (2001). Synthesizing Life. Nature 409:387-390.