DOI: http://dx.doi.org/10.7551/978-0-262-32621-6-ch131

An Artificial Chemistry for the 'Lipid World' Scenario

Barak Shenhav

Department of Software Engineering, Afeka - Academic College of Engineering, Tel aviv, Israel
baraks@afeka.ac.il

Abstract

'Metabolism first' scenarios are amply suggested as an alternative or precursor for the emergence of the 'RNA world' that may have prevailed before contemporary life as we know it. The 'Lipid world' scenario advocates 'metabolism first'. The commonly used computational model for the scenario (the GARD model) emphasizes that reaction catalysis is not a binary phenomenon, but rather a matter of magnitude. Simulations of the model revealed lack (or at most very low capacity) to undergo natural selection. Here, it is proposed that this cavity emerges from the underlying distribution of catalytic values typically employed by the model. In particular, it is shown that the propensity of having many (or even few) 'viable' autocatalytic cores in a GARD system is rather slim. The robustness of GARD cores to parasitic periphery is pinpointed. As a conclusion, it is suggested that a 'Lipid world' based on a simple artificial chemistry may harbor many cores, possibly infinite, facilitating evolutionary process.

Introduction

The 'Lipid world' scenario [Segre et al. 2001a] suggests that life may have emerged from non-covalent molecular assemblies. Such aggregates may be micelles, bilayers or vesicles that spontaneously rise when amphiphilic molecules (lipids) are introduced to an aqueous environment. Chemical studies and analysis of the scenario go back many decades [Oparin, 1957]. Theoretical studies of this scenario utilize the Graded Autocatalysis Replication Domain (GARD) model [Segre et al. 2000], which is briefly described in the following section.

Based on GARD simulations, it was postulated that self-sustaining autocatalytic networks lack evolvability [Vasas et al. 2010]. Farther studies on other models of autocatalytic networks [Kauffman, 1993] suggested that while most models fail to portray evolutionary capacity altogether, some may have such capacity, though more research is required in that realm [Vasas et al. 2012]. A detailed study, by the group that originated the GARD model, suggests that mutual catalysis is fundamental for evolvability [Markovitch and Lancet, 2012]. However, this later study (cf. figure 6 therein) show that a typical GARD system have a single compotype (conceptually a compotype may be thought of as an attractor or a quasi-stationary state in the network dynamics), and no simulation (out of 10,000) showed more than six compotypes.

Modeling the 'Lipid World'

The basic GARD model (dealt with hereafter) is thoroughly described in numerous publications (see ool.weizmann.ac.il for a partial list including several other variants of the model). Follows, a concise description of the model (adopted from [Vasas et al. 2010]):

> It (the model – the author) involves discrete stochastic changes in noncovalent assemblies dictated by the differential equations
> N_G is the molecular repertoire of environmentally available
> $$\frac{dn_i}{dt} = (\rho_i k_i N - k_{-i} n_i)\left(1 + \frac{1}{N}\sum_{j=1}^{j=N_G} \beta_{ij} n_j\right) \quad i = 1, 2, \ldots, N_G$$
> prebiotic compounds; ρ_i is the external concentration of molecular species i; $k_i = 10^{-2}\text{sec}^{-1}$ and $k_{-i} = 10^{-5}\text{sec}^{-1}$ are uncatalyzed forward and backward rate constants assumed to be equal for all molecules for simplicity [they differ in their mutual rate enhancement properties]; N ($N<N_G$) is the assembly size given by $N=\Sigma n_i$, with n_i indicating the count of molecular species i (within the assembly – the author) and β_{ij} is an element of the $N_G \times N_G$ positive matrix that defines the network of mutually catalytic interactions governed by a statistical formalism (see below)

The course of the composition of an assembly is governed by the equations above. It was shown that if assembly expansion is not disrupted, the assembly reaches an asymptotic steady-state composition [Segre et al. 2001b]. A more complex dynamics rise when a GARD assembly goes through a growth-fission process. Fission of molecular assemblies is likely to occur as the assembly grows larger. In GARD this was simplified proposing that as the assembly size (N) doubles the assembly breaks into two daughter assemblies, where each molecule in the parent assembly has 50% probability to go to either daughter assembly. Adding fission process keeps the assembly out of equilibrium. Fission is necessary (though not sufficient) to allow the exsistence of several quasistationary compositions (composomes) in a GARD system.

The matrix β defines the amount of mutual catalysis exerted on the join/leave reactions depicted by the kinetic equations. The elements β_{ij} of the matrix are drawn from a log-normal distribution with parameters $\mu=-4$ and $\sigma=4$. The log-normal distribution has a long tail, allowing seldom high values, often many orders of magnitude above the background average. The log-normal distribution is an approximation of the Receptor Affinity Distribution [Lancet, 1993] modified for catalytic rate enhancement.

Results

The values in the β matrix are generated independently. Thus, the distribution of values in the j^{th} row of β (that is the catalysis exerted by molecular species j on all N_G molecular species) is similar to the distribution of the values in whole β. The probability that the largest value in the j^{th} row will occur in the j^{th} column (i.e. on the diagonal of β) is $1/N_G$. The probability that for all N_G rows the largest value (of each row) do not occur on the diagonal is therefore:

$$\left(1-1/N_G\right)^{N_G}$$

which is approximately 1/e or 36.79%. Hence, most GARD systems (63.21%) are likely to hold at least one auto-catalytic molecular species. We shall denote these systems as Class-1.

Now, let us consider other classes of systems. Class-2 systems harbor a core (cf. Vasas et al. 2012) of two molecular species (if molecular species A and B are the core, the strongest catalysis of A is on the reaction that generates B and the strongest catalysis of B is on the reaction that generates A), but do not hold an auto-catalyst. Class-3 systems are neither in Class-1 nor in Class-2. Similar to analysis in the previous paragraph, the probability for a system to be in Class-3 is:

$$\left(1-1/N_G-N_G\cdot 1/N_G\cdot 1/N_G\right)^{N_G}$$

which is approximately 1/e squared or 13.53%. This leads to the conclusion that the probability of a system to be in Class-2 is 23.25%. (36.79%-13.53%).

It should be noted that given the characteristics of the distribution that generates the entries of β the largest value in a row is very likely to be larger than the sum of all other entries in the row. Consequently, the effect of parasitic periphery on the sustainability of the core diminishes. This strength of the GARD model is generally ignored elsewhere, and should be further analyzed.

The detailed outcome of a specific GARD system is determined by the holistic contribution of the catalysis. Yet, the dynamic corresponds, usually, to the classification of the system:

- Class-1 systems are generally governed by the strongest auto-catalyst that out-competes all other auto-catalysts. The governing compotype may "vanish" for some periods where fabricated-composomes (a very strong catalyst that "pumps" its substrate on expense of other species as long as the catalyst is within the assembly). Conceivably, other cores, if exist, may compete with the governing auto-catalyst. Yet, as larger cores lack robustness (compared to the auto-catalyst) they rarely prevail.
- Class-2 systems show, typically, a mixture of core based composomes (if several small cores exist in the network) and fabricated-composomes as well as some periods of random compositional drift.
- Class-3 systems are generally in a drift state as with seldom appearances of fabricated-composomes. The large core generally cannot "squeeze" into the assembly, in line of the analysis briefly portrayed earlier [Shenhav et al. 2004]

The analysis and study of GARD classes is in progress.

Discussion

The theoretical analysis above is in line with the empirical results that GARD systems commonly show a single or at most few compotypes. Nonetheless, based on the analysis, though requiring additional effort, GARD systems may be classified according to features in their underlying network (which may be a rough analogy to genotype) rather than according to the dynamics the system manifests (phenotype).

GARD was severely criticized for its lack of evovability, Here, it is suggested that this deficiency is due to the nature of the underlying distribution of catalysis commonly used in the model. It is likely that engaging with an alternative structure for the β matrix, for example a block-diagonal, should result to multiple 'viable' cores, with possible better evolvable capacity. Amending the underlying chemistry resembles suggestions by others [e.g. Giri and Jain, 2012] regarding "fixing" a delinquent model. While the chemical plausibility of such matrix structure may be questioned, pursuing this kind of artificial chemistry seems promising.

The extension of the analysis to other GARD variants [Shenhav et al. 2007] is far from straight-forward.

References

Giri V. and Jain (2012). The origin of large molecules in primordial autocatalytic reaction networks. *PLoS ONE*, 7(1): e29546

Kauffman S. A. (1993). *Origins of Ord*er. Oxford University Press, New York, NY.

Lancet D., Sadovsky E. and Seidemann E. (1993). Probability model for molecular recognition in biological receptor repertoires - Significance to the olfactory system. *Proceedings of the National Academy of Sciences of the United States of America*, 90(8):3715–3719.

Markovitch O. and Lancet D. (2012). Excess mutual catalysis is required for effective evolvability. *Artificial Life*, 18(3):243–266.

Oparin, A. I. (1957). *The Origin of Life on the Earth*. Oliver and Boyd, London

Segre D., Ben-Eli D. and Lancet D. (2000). Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. *Proceedings of the National Academy of Sciences of the United States of America*, 97(8):4112–4117.

Segre D., Ben-Eli D., Deamer D. and Lancet D. (2001a). The lipid world. *Origins of life and evolution of the biosphere*, 31:119–145.

Segre D., Shenhav B., Kafri R. and Lancet D. (2001b). The molecular roots of compositional inheritance. *Journal of Theoretical Biology*, 213:481–491.

Shenhav B., Kafri R. and Lancet D. (2004). Graded artificial chemistry in restricted boundaries. In Pollack J., Bedau M.A., Husbands P., Watson R.A. and Ikegami T., editors, *Artificial Life IX*, pages 501–506. MIT Press, Cambridge, MA.

Shenhav B., Oz A. and Lancet D. (2007). Coevolution of compositional protocells and their environment. *Philosophical Transactions of the Royal Society B*, 362: 1813–1819.

Vasas V., Szathmary E. and Santos M. (2010). Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *Proceedings of the National Academy of Sciences*,107(4):1470-1475.

Vasas V., Fernando C., Santos M., Kauffman S. and Szathmary E. (2012). Evolution before genes. *Biology Direct*,7:1