THE LIPID WORLD: FROM CATALYTIC AND INFORMATIONAL HEADGROUPS TO MICELLE REPLICATION AND EVOLUTION WITHOUT NUCLEIC ACIDS

ARREN BAR-EVEN, BARAK SHENHAV, RAN KAFRI AND DORON LANCET
Department of Molecular Genetics and the Crown Human Genome Center, the Weizmann Institute of Science, Rehovot 76100, Israel

1. Lipid World And The GARD Model

A widespread notion is that life arose from a single molecular replicator, probably a self-copying polynucleotide, in an RNA World (Joyce, 2002). We have proposed an alternative Lipid World scenario as an early evolutionary step in the emergence of cellular life on Earth (Segre *et al.*, 2001). This concept combines the potential chemical activities of lipids and other amphiphiles, with their capacity to undergo spontaneous self-organization into supramolecular structures, such as micelles and bilayers. In quantitative, chemically-realistic computer simulations of our Graded Autocatalysis Replication Domain (GARD) model (Segre *et al.*, 1998), we have shown that prebiotic molecular networks, potentially existing within assemblies of lipid-like molecules, manifest a behavior similar to self reproduction or self-replication.

Because amphiphile assemblies may form readily and spontaneously under prebiotic conditions (Deamer, 1985), the Lipid World scenario may represent an intermediate "mesobiotic" phase, bridging an a-biotic random collection of organic molecules with a biotic protocell that contains long biopolymers, as well as more intricate information storage, catalysis and replication (Shenhav *et al.*, 2003). In our model, lipid-like amphiphiles may possess a very large variety of chemical structures, including head-groups that resemble amino-acids or nucleotides. Catalysis is proposed to be exerted by such diverse chemical moieties, enhancing amphiphile exchange rates as well the formation of more complex head-groups with similarity to peptides or oligonucleotides. In a more recent version of our model (Polymer GARD or P-GARD), a path is delineated for the gradually increased dependence on linear molecular sequences (Shenhav *et al.*, in press).

A most central notion in our P-GARD model is that life began with monomers only, and later proceeded to evolve biopolymers. This is in strict contrast to the widespread view, epitomized by the RNA World concept, whereby the abiotic formation of biopolymers is a prerequisite for life. In other words, RNA World or protein World proponents assert that a molecular entities devoid of biopolymers cannot possess the basic attributes of life. This credo derives from the premise that a living entity has to contain information and transmit it to progeny via replication, and that such a sequence

of events cannot occur without long informational molecules that form similes by base pairing, i.e. without nucleic acids.

The unorthodoxy of the Lipid World notion thus goes beyond the question of what chemical entities formed the first replicators – nucleotides, amino acids or lipids. In fact, Lipid World in its most general sense does not prescribe a particular chemistry, and only suggests that the first events en route to life were mediated by molecules capable of readily forming non-covalent aggregates. This allows a much wider diversity of compounds to take part in the first events leading to life, hence generates a more plausible scenario for life's beginning in a highly heterogeneous primordial "soup".

In the basic model of monomer GARD, quasi-stationary assembly compositions (composomes) are shown to survive many splits, before giving rise to other quasi-stationary state or being scattered into "random drift" composition (Segre *et al.*, 2000). Two major ideas underlie the GARD model. First is the notion of compositional information, namely that information can be stored in the distribution of compounds within the entity, rather by a sequence embodied within each compound. In other words, information is not a characteristic of an individual molecule, but of the system as a whole. Second is the capacity of that information to be transmitted along generations with acceptable fidelity. The transition between one composomal state and another can be viewed as an evolution-like process, in which one organism gives rise to another due to stochastic progression.

2. Fitness Parameters And Populations of Assemblies

Previously, we have mainly dealt with individual GARD assemblies, and at each split, one was randomly selected for further scrutiny. However, evolution is mostly about populations of organisms, displaying "ecological" relationships such as competition. We therefore have initiated a study of the properties of GARD assembly populations. Because of computing power limitations it is imperative to generate a phenomenological description of every composome in the realm of a higher level of analytical hierarchy. This is done by describing every composome with three kind of emergent property parameters (Shenhav *et al.*, in press). First is growth time, T_i, the time elapsing between assembly divisions, while the assembly is in a specific composomal state i. Second is the survival perameter (S_i), namely the probability of an assembly to preserve its composomal state following a split. Last is the emergence likelihood (E_i), which is the probability of entering a composomal state from another one or from random drift (a state in which the assembly possesses no composomal state). In a given population of assemblies and for specific environment attributes, these fitness parameters will provide a good approximation of the composomes population dynamics. In fact, it is possible to display this dynamics in a set of linear differential equations, using an "ecological" matrix that can be constructed using the three fitness parameters.

Our simulations showed that in most cases there is a dominant composome, in term of the S_i and E_i, though other composomes might be much faster (small T_i) and therefore can be present in the population in a significant fraction (Fig. 1). Moreover, repeated simulations demonstrated that if the system is left to evolve without any constraint, the population distribution reaches a steady state, that could be interpreted as a stable "eco-system". In any given set of parameters this state is unique, regardless of the initial

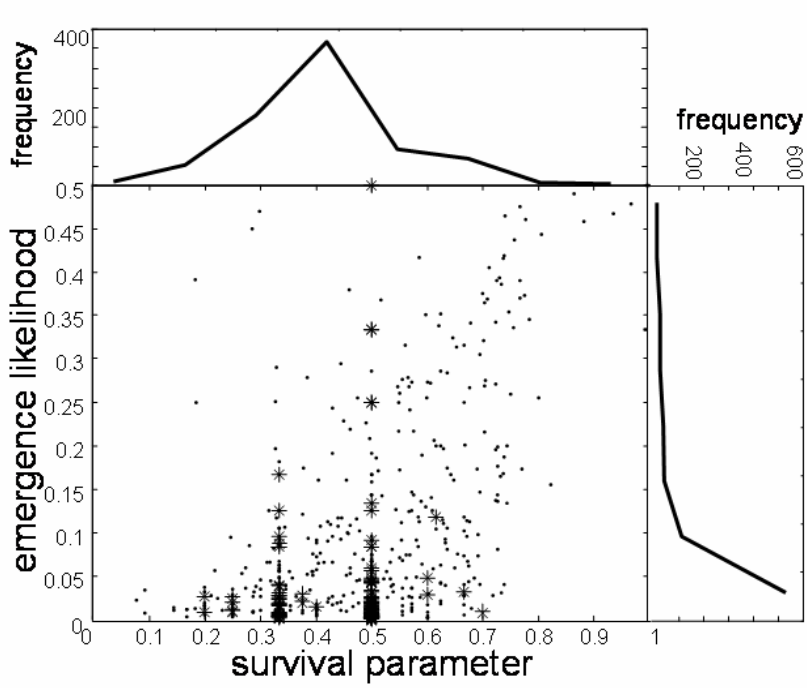

Figure 1. Correlation between two of the fitness parameters. Dots represents slow composome (large T_i) and asterisks denote the top 20% fastest composomes (low T_i). Overall, there are 786 composomes, obtained from 100 Dimer-GARD simulations with different catalytic enhancement matrices, all drawn from the same catalytic values distribution (Segre *et al.,* 2000). The survival parameter (S_i) values behave roughly as a normal distribution, while the emergence likelihood (E_i) values obey an exponential-like distribution. There is a rough linear correlation between those two fitness parameters. It is interesting to note that the dominant composomes (having both large Si and Ei) tend to be slow, allowing the fast growing ones, located mostly in the middle bottom of the graph, to successfully compete, resulting with their significant representation in the population.

population seeded. Imposing population constraints, such as constant population, prevents the system from converging on one state for long periods, though quasi-stationary population distribution can survive for substantial time before giving rise to an other. Interestingly, in the linear approximation of the "eco-system" dynamics that we have begun to pursue, the dominant population distribution corresponds to the all-positive eigen-vector of the “ecological matrix”, in analogy to the dominant GARD composome being related to the all-positive eigen-vector of the catalytic-enhancements beta matrix (Segre *et al.*, 2000).

3. Towards Increasing Complexity – Synthesis of Polymers

Our Monomer-GARD simulations (Segre *et al.*, 2000) have a limited capability of creating metabolism-like systems with relatively high complexity, which is one of the basic features of evolution. In order to generate possible trajectories towards ever increasing complexity, we have introduced into our simulations a more elaborate form

of metabolism, namely the capability of monomer to join each other to form oligomers. The basic interactions logic remains unchanged, and the only additional type of allowed interaction among molecules is the catalytic enhancement upon the oligomerization and bond-breaking rates.

First we have used monovalent dimer and therefore restricted oligomerization to dimers only (Shenhav *et al.*, in press). Fig. 1 presents the distribution of the fitness parameters calculated for the 786 composomes obtained in one hundred such Dimer-GARD simulations with different catalytic enhancement matrices.

Introducing divalent monomers, capable of generating polymer with unlimited length, highlights a central problem of such simulations, namely how to relate the dependence of the catalytic capacity of longer molecules to the catalytic capacity of the monomers composing it. At the one end are models in which the catalytic enhancement of a molecule is the sum of those of its monomers, and therefore the oligomerization contributes minimally to the overall catalysis in the system. An alternative scenario is that the catalytic characteristics of an oligomer are unrelated to those of its monomers, i.e. potencies are randomly assigned. The models of Dyson (Dyson, 1999) and Kaufmann (Kauffman, 1993) are good examples of these two options, respectively. At another extreme, oligomer-induced catalysis is non-linearly related to those of the constituent monomers, e.g. a product thereof, underlying synergism. Our preliminary results show that in a purely linear model, stable composomes do not form readily. The desirable model should compromise two opposing trends – the ability to sustain stable composomes while giving enough space for the evolution of more complex quasi-stationary states, in the course of time.

Our main research objective is to define a basic set of interactions rules so that primitive genetic memory will become an emergent property, alongside with an increased complexity of the system. We anticipate that this could be achieved by a carefully designed set of oligomerization rules. It is not impossible that without synergism the system will not develop significance memory and complexity. Implementation of extended oligomer-based models could help bridge the gap between simple and close-ended systems to more elaborate ones that are essentially open-ended and therefore can serve as a true platform for the origin of life.

4. References

Deamer, D.W. (1985) Boundary Structures are Formed by Organic Components of the Murchison Carbonaceous Chondrite. Nature **317**, 792–794.

Dyson, F. (1999) *Origins of life*, Cambridge University Press, Cambridge.

Joyce, G.F. (2002) The antiquity of RNA-based evolution, Nature **418**, 214-221.

Kauffman, S. (1993) *The origin of order*, Oxford University Press.

Segre, D., Lancet, D., Kedem, O., Pilpel, Y. (2001) Graded autocatalysis replication domain (GARD): Kinetic analysis of self-replication in mutually catalytic sets. Orig. Life Evo. Biosph. **28**, 501-514

Segre, D., Ben-Eli, D. and Lancet, D. (2000) Compositional genomes: Prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc. Natl. Acad. Sci. USA **97**, 4112-4117.

Segre, D., Ben-Eli, D., Deamer D.W. and Lancet, D. (2001) The lipid world. Orig Life Evol. Biosph. **31**, 119-145

Shenhav, B., Segre D. and Lancet D. (2003), Mesobiotic emergence: Molecular and ensemble complexity in early evolution. Adv. Complex Syst. **6**, 15-35.

Shenhav, B., Bar-Even, A., Kafri, R. and Lancet, D., Polymer GARD: simulation of covalent bond formation in reproducing molecular assemblies. Orig Life Evol. Biosph, in press.