DOI: http://dx.doi.org/10.7551/978-0-262-31709-2-ch020

Prebiotic Evolution of Molecular Assemblies: From Molecules to Ecology

Markovitch Omer[1] and Lancet Doron[1]

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel
omermar@gmail.com

Abstract

Present life portrays a two-tier phenomenology: molecules compose supramolecular structures, such as cells or organisms, which in turn portray population behaviors, including selection, evolution and ecological dynamics. Prebiotic models have often focused on evolution in populations of self-replicating supramolecules, without explicitly invoking the intermediate molecular-to-supramolecular stage. We explore a prebiotic model that allows one to relate parameters of chemical interaction networks within molecular assemblies to emergent ecological and evolutionary properties in populations of such assemblies. We use the graded autocatalysis replication domain (GARD) model, which simulates the network dynamics of amphipile-containing molecular assemblies, and exhibits quasi-stationary compositional states termed compotypes. These grow by catalyzed accretion, divide and propagate their compositional information to progeny in a replication-like manner. The model allows us to ask how molecular network parameters influence assembly evolution and population ecology, analyzable by a multi species logistic (r-K) model for population ecology (Lotka-Volterra competition model). We found that compotypes with a larger intrinsic molecular repertoire show a higher intrinsic growth (r) and lower carrying capacity (K), as well as lower replication fidelity. This supports a prebiotic scenario initiated by fast-replicating assemblies with a high molecular diversity, evolving into more faithful replicators with narrower molecular repertoires. A main difference from classical ecology is that in GARD species inter convert into each other rather than consume each other or compete on resources, thus representing 'fast forward' of speciation.

Introduction

The path from organic mixtures (i.e., the primeval soup) to reproducing life-like protocells no doubt required the emergence of replicating systems capable of undergoing Darwinian evolution. Therefore, uncovering how such entities emerged in early niches will greatly contribute to our understanding of life's origin and can potentially allow one to design novel experiments. The GARD model (Segre, Ben-Eli,Lancet 2000) in the realm of the lipid world scenario (Segre, Ben-Eli, Deamer et al. 2001) offers one possible route for such pursuit. In this framework, non-covalent assemblies of amphiphiles, such as lipid micelles or vesicles are studied. These store information in the form of nonrandom molecular compositions, which is passed to progeny via homeostatic growth accompanied by fission. The model quantitatively describes the details of such a process (Segre et al. 2000). It is based on a directed catalytic network (termed β), whose nodes and edges respectively represent molecular types and catalytic rate enhancements. Importantly, the system is kept away from thermodynamic equilibrium by assembly fission, which produces two progeny assemblies. Key in GARD dynamics are *composomes*, replication-prone quasi-stationary states. A group of composomes, gleaned by clustering, is termed *compotype*, and may be regarded as species in the framework of lipid world and GARD. Indeed, such GARD species were recently shown to display a significant measure of Darwinian evolution (Markovitch and Lancet 2012), in disagreement with a report (Vasas, Szathmáry,Santos 2010) criticizing this notion on the basis of testing random compositions and with no statistical rigor.

Simulations

The GARD10 MATALB code was employed for all simulations, using parameter values identical to those employed previously (Markovitch and Lancet 2012). The dynamics of compositional assemblies in a reactor under constant population conditions were examined. The reactor is seeded with 1,000 random compositions which are allowed to simultaneously grow based on their idiosyncratic kinetic parameters, and undergo fission when reaching a predefined maximal size. The pre-fission composition of each assembly is assessed as belonging to one of the compotypes characterizing the specific β or to "drift". C_i marks the fractional number of assemblies belonging to compotype i (out of 1,000). Each simulation is performed for 50,000 split events in the reactor, typically sufficient to reach steady state in compotype frequencies. For statistical rigor, 1,000 such simulations were performed, each with a different β whose edges are randomly drawn from a lognormal distribution (Segre, Shenhav, Kafri et al. 2001).

GARD Population Dynamics

Different simulations showed widely different dynamic behaviors, such as non-trivial "takeover" of a fast-rising compotype by a slower one (Fig. 1). Such dynamics is typical of natural ecosystems that harbor multiple species with competition or predator-prey relationships. The results were analyzed by a multi species logistic equation (May 1974; Gabriel, Saucy,Bersier 2005) (Fig. 1 legend). The GARD model thus affords a unique opportunity to directly relate molecular parameters to ecological behavior, bypassing the organismal complexity that usually bridges the two. As an example, the relationship between a compotype's molecular

diversity and two central quantitative ecological parameters are portrayed here.

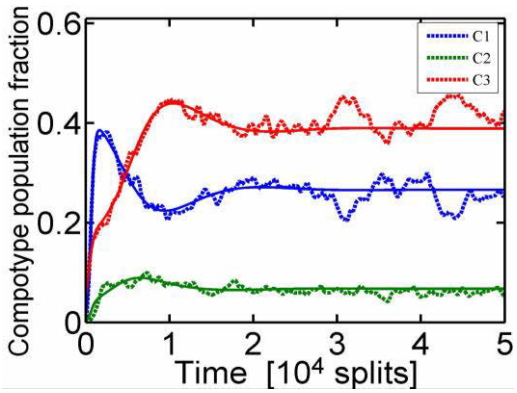

Fig. 1: An example of population dynamics. Broken lines represent the three compotype species found in this simulation and solid lines are a fit to the logistic growth: $dC_i/dt=r_i*C_i*[K_i-C_i-\sum(\alpha_{ij}*C_j)]/K_i$, where C_i is the population-fraction of compotype i in the population at time t. Fitted parameter values are: $r_{1..3}$= 5e-3, 4e-3, 2e-3; $K_{1..3}$= 0.56, 0.49, 0.72; $C(t=0)_{1..3}$= 0.019, 0.003, 0.055; α_{12}=1.5, α_{13}=0.05, α_{21}=0.8, α_{23}=0.56, α_{31}=1.3, α_{32}=0.

Compotype Molecular Diversity

The chemistry to ecology transition was done by analyzing populations of GARD assemblies through the scope of multi species logistic growth. In this analysis, each compotype species i is characterized by two basic parameters: the intrinsic growth rate (r_i) and the carrying capacity (K_i).

The simulations result in the generation of about 1,500 compotypes in 1,000 chemical niches (β networks). Compotypes are defined at their molecular level by N_{mol}, the size of the intrinsic molecular repertoire of the compotype. In a simulation with N_G molecular types, $N_{mol}<N_G$ represents the subset of molecule types present as a result of the intermolecular catalytic interactions in β. It is found that K values are inversely correlated with N_{mol}. In contrast, r values show a weak positive correlation (Fig. 2). Thus, in the absence of competition, the time-dependent prevalence of compotypes with large N_{mol} will show a steep ascent with a relatively low plateau, while those with low N_{mol} will show a slower ascent but can potentially reach a higher plateau.

Of note, an analog trend to the increase of the intrinsic growth rate with molecular diversity was observed in experimental data for 113 Bacteria (Freilich, Kreimer, Borenstein et al. 2009), whereby a negative correlation between measured doubling time and metabolic network size was found. The results might advocate for a prebiotic scenario initiated by fast-replicating assemblies with a high molecular diversity, evolving into more faithful replicators with narrower molecular repertoires. This is not unlike the transition from prebiotic "random chemistry" to the relatively restricted repertoire of small molecules (monomers) seen in present-day living cells (Segre et al. 2000).

Compotype Replication Fidelity

In the absence of competition, i.e. when examining simulations exhibiting only one compotype species, a

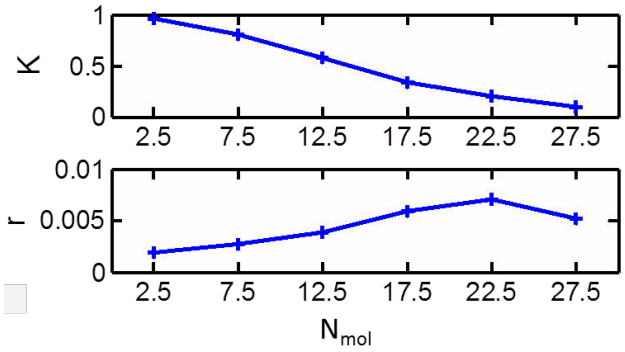

Fig. 2: Compotypes dependence of carrying capacity (K) and intrinsic growth rate (r) on the size of the intrinsic molecular repertoire of compotypes (N_{mol}). Data is binned.

compotypes' K values typically does not reach the upper limit of 1.0 (mean K=0.54±0.35 for such simulations). K represents the maximal number of individuals that may be sustained in an environmental niche. In the original Verhulst formalism, death was introduced by as a potential solution to the Malthusian exponential growth, and later in the r-K formalism K=birth/death (Gabriel et al. 2005). In GARD, a similar interpretation of K pertains, whereby a positive correlation between K and replication fidelity (F_{rep}) is observed: $K=6.2*F_{rep}-5.35$ with $R^2=0.63$. F_{rep} measures the average degree of compositional similarity between a compotype assembly to its fully grown progeny. Thus, unfaithful replication means that the fully grown progeny has lost its compotype state and is considered drift, somewhat comparable to death.

Acknowledgements

We thank Paul Higgs, Raphael Zidovezki and Natalio Krasnogor for discussions. This work is partly supported by EU-FP7 project MATCHIT and by the Crown Human Genome Center at the Weizmann Institute of Science.

References

Freilich S, Kreimer A *et al.* (2009) Metabolic-network-driven analysis of bacterial ecological strategies. *Genome Biol* 10 (6).

Gabriel JR, Saucy F *et al.* (2005) Paradoxes in the logistic equation? *Ecol Model* 185 (1):147-151.

Markovitch O, Lancet D (2012) Excess Mutual Catalysis Is Required for Effective Evolvability. *Artif Life* 18 (3):243-266.

May RM (1974) Biological Populations with Nonoverlapping Generations - Stable Points, Stable Cycles, and Chaos. *Science* 186 (4164):645-647.

Segre D, Ben-Eli D *et al.* (2001) The lipid world. *Origins of Life and Evolution of the Biosphere* 31 (1-2):119-145.

Segre D, Ben-Eli D *et al.* (2000) Compositional genomes: Prebiotic information transfer in mutually catalytic noncovalent assemblies. *P Natl Acad Sci USA* 97 (8):4112-4117.

Segre D, Shenhav B *et al.* (2001) The molecular roots of compositional inheritance. *Journal of Theoretical Biology* 213 (3):481-491.

Vasas V, Szathmáry E *et al.* (2010) Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *P Natl Acad Sci USA* 107 (4):1470-1475.