Journal of Theoretical Biology 381 (2015) 29–38

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Primordial evolvability: Impasses and challenges

Vera Vasas [a], Chrisantha Fernando [b], András Szilágyi [c,d], István Zachár [d], Mauro Santos [e,*], Eörs Szathmáry [c,d]

[a] School of Biological and Chemical Sciences (SBCS), Queen Mary University of London, Mile End Road, London, UK
[b] School of Electronic Engineering and Computer Science (EECS), Queen Mary University of London, Mile End Road, London, UK
[c] Parmenides Center for the Conceptual Foundations of Science, Kirchplatz 1 Pullach, D-82049 Munich, Germany
[d] Biological Institute and MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Eötvös University, Budapest, Hungary
[e] Departament de Genètica i de Microbiologia; Grup de Genòmica, Bioinformàtica i Biologia Evolutiva (GGBE), Universitat Autonòma de Barcelona, Bellaterra, Barcelona 08193, Spain

HIGHLIGHTS

- Incorporation-type chemistry is unfavourable for evolvability.
- Lognormal distribution of catalytic factors hinders autocatalysis.
- The GARD model is dominated by strong non-autocatalytic components.
- Real chemical reactions that make or break covalent bonds are necessary for appreciable evolvability.
- Limited, but substantial, heredity is needed for evolvability without templates.

ARTICLE INFO

Article history:
Received 10 March 2015
Received in revised form
22 June 2015
Accepted 24 June 2015
Available online 9 July 2015

Keywords:
Origin of life
Chemical evolution
Autocatalysis
Catalytic reaction networks
Metabolism-first theories
Collectively autocatalytic sets
Evolvability
GARD model

ABSTRACT

While it is generally agreed that some kind of replicating non-living compounds were the precursors of life, there is much debate over their possible chemical nature. Metabolism-first approaches propose that mutually catalytic sets of simple organic molecules could be capable of self-replication and rudimentary chemical evolution. In particular, the graded autocatalysis replication domain (GARD) model, depicting assemblies of amphiphilic molecules, has received considerable interest. The system propagates compositional information across generations and is suggested to be a target of natural selection. However, evolutionary simulations indicate that the system lacks selectability (i.e. selection has negligible effect on the equilibrium concentrations). We elaborate on the lessons learnt from the example of the GARD model and, more widely, on the issue of evolvability, and discuss the implications for similar metabolism-first scenarios. We found that simple incorporation-type chemistry based on non-covalent bonds, as assumed in GARD, is unlikely to result in alternative autocatalytic cycles when catalytic interactions are randomly distributed. An even more serious problem stems from the lognormal distribution of catalytic factors, causing inherent kinetic instability of such loops, due to the dominance of efficiently catalyzed components that fail to return catalytic aid. Accordingly, the dynamics of the GARD model is dominated by strongly catalytic, but not auto-catalytic, molecules. Without effective autocatalysis, stable hereditary propagation is not possible. Many repetitions and different scaling of the model come to no rescue. Despite all attempts to show the contrary, the GARD model is not evolvable, in contrast to reflexively autocatalytic networks, complemented by rare uncatalyzed reactions and compartmentation. The latter networks, resting on the creation and breakage of chemical bonds, can generate novel ('mutant') autocatalytic loops from a given set of environmentally available compounds. Real chemical reactions that make or break covalent bonds, rather than mere incorporation of components, are necessary for open-ended evolvability. The issue of whether or not several concrete chemical systems (rather than singular curiosities) could realize reflexively autocatalytic macromolecular networks will ultimately determine the relevance of metabolism-first approaches to the origin of life, as stepping stones towards true open-endedness that requires the combination of rich combinatorial chemistry controlled by information stored in template replicators.

* Corresponding author.
E-mail addresses: vvasas@yahoo.com (V. Vasas), ctf20@eecs.qmul.ac.uk (C. Fernando), and.szilagyi@gmail.com (A. Szilágyi), replicatorzed@gmail.com (I. Zachár), mauro.santos@uab.es (M. Santos), szathmary.eors@gmail.com (E. Szathmáry).

http://dx.doi.org/10.1016/j.jtbi.2015.06.047
0022-5193/

1. Introduction

In an influential paper on the units of selection problem in biology, Lewontin (1970) set out the three principles that embody Darwin's scheme of evolution as it was seen by evolutionary biologists: phenotypic variation, fitness differences and heritability of fitness. Some years later Maynard Smith (1983, 1987) apparently rephrased this account and offered the principles of variation, multiplication and heredity as the basic pillars of Darwin's theory of evolution by natural selection: "…. if there is a population of entities with multiplication, variation, and heredity, and if some of the variations alter the probability of multiplying, then the population will evolve. Further, it will evolve so that the entities come to have adaptations". Albeit superficially similar, Lewontin's and Maynard Smith's accounts are deemed fundamentally different according to Griesemer (2000). However, for our present purposes it would suffice to say that these principles are general and no particular mechanism of inheritance is assumed; only a statistical correlation between parent and offspring. Nevertheless, even a cursory reading of the literature dealing with the units of selection shows clearly that genetic inheritance is taken to mean the transfer of digitally encoded information. Whether this should be a necessary condition for units of evolution, or simply reflects a conceptual bias arisen from what happens in present-day cells, is a fundamental question to be solved in systems chemistry dealing with the "conjunction of supramolecular and prebiotic chemistry with theoretical biology and complex systems research addressing problems relating to the origins and synthesis of life" (von Kiedrowski et al., 2010).

The importance of defining what a unit of Darwinian evolution is becomes critically relevant in the context of the emergence of life because it outlines a sharp divide between abiogenesis (prior chemical evolution) and the very origin of evolvability; namely, the capacity of a system to experience adaptive evolution. This obviously happens if genetic inheritance is due to template-directed replication of nucleic acids, as exemplified by the extraordinary complexity and intricate organization of living things. The crucial question is: can adaptive evolution happen when information transfer is non-digital? More specifically, are those systems where the transfer of information arises as a parent-offspring correlation in molecular composition evolvable?

2. The GARD model: Basics and challenges

Doron Lancet and collaborators (Segré et al., 1998, 2000) published a promising model of a potentially evolvable, template-free prebiotic system, called GARD (Graded Autocatalytic Replication Domain; Fig. 1). The model is an example of the lipid world scenario (Segré et al., 2001a). Essentially, the authors envisage a multitude of lipid assemblies (micelles or vesicles) that spontaneously grow due to the incorporation of lipid monomers from the environment.

The characteristic equations that describe GARD assemblies are (Segré et al., 2000):

$$\frac{dn_i}{dt} = F_i(\boldsymbol{\eta}^{\mathbf{G}}) = (\rho_i k_i N - k_{-i} n_i)\left(1 + \frac{1}{N}\sum_{j=1}^{j=N_G} \beta_{ij} n_j\right), \quad i = 1, 2, \cdots, N_G \tag{1}$$

where $\boldsymbol{\eta}^{\mathbf{G}}$ is an N_G – long vector;N_G is the molecular repertoire of environmentally available prebiotic compounds; ρ_i is the external concentration of molecular species i; $k_i = 10^{-2}\ s^{-1}$ and $k_{-i} = 10^{-5}\ s^{-1}$ are (arbitrary) uncatalyzed forward and backward rate constants assumed to be equal for all molecules for simplicity. $N < N_G$ is the assembly size given by $N = \sum_{i=1}^{N_G} n_i$, with n_i indicating the count of molecular species i; that is, the internal molecular counts of vector $\boldsymbol{\eta}^{\mathbf{G}}$ are $n_1, n_2, \ldots, n_{N_G}$ (the constraint $N < N_G$ is because when $N > N_G$ the information transmission is trivial for lack of compositional variation; Segré et al., 2001b). The crucial parameters are the β_{ij} values, the elements of the $N_G \times N_G$ positive matrix $\boldsymbol{\beta}$ that defines the network of mutually catalytic interactions. The diagonal elements describe autocatalytic molecules and the off-diagonal elements cross-catalytic ones. For large N_G values, this will result in a complex mutually catalytic network, represented by the $\boldsymbol{\beta}$ matrix. The elements of this matrix are drawn from a log-normal distribution, as an approximation of the receptor affinity distribution modified for catalytic rate enhancement (Lancet et al., 1993), but note that there is no particular physical-chemical constraint that imposes such distribution. The matrix may or may not have diagonal elements (direct autocatalysis) and naturally it is not symmetrical. Thus, a central assumption of the GARD model is that molecules already present within an assembly may enhance the rate of joining and leaving of new molecular species, and that this feature is specific to each individual molecular species. Even if all diagonal elements are zero, this molecular system as a whole is collectively autocatalytic *sensu* Kauffman (1986) and thus qualifies as an ensemble replicator (Szathmáry, 2000). Collective autocatalysis arises when components of a system are not necessarily autocatalytic, but they catalyse each other's formation/entry in such a way that formation of every member in the set is catalyzed by the formation of at least one other member in the set.

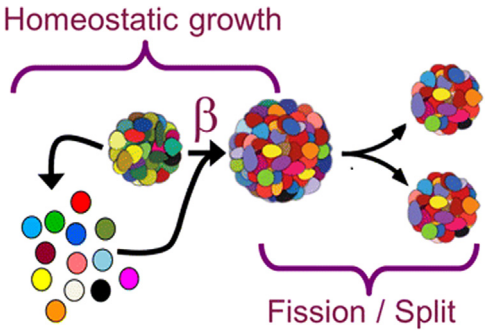

Fig. 1. A general scheme of the GARD model. The environment contains a large repertoire of environmentally available amphiphiles. At any time a GARD assembly contains a subrepertoire of molecular types. The assembly grows by the accretion of amphiphiles, which is dictated by the positive matrix β that defines the network of mutually catalytic interactions governed by a statistical formalism. Once the assembly has reached a predefined maximal size, a binary fission occurs and the growth cycle begins again. From Gross et al. (2014).

When the GARD assembly is assumed to go through a growth-splitting process, a non-trivial behaviour emerges. Fission is assumed to happen when the size of the assembly (N) reaches a threshold value. This process, imposed by surface tension or turbulence, serves as an external free energy input and keeps the assemblies out of thermodynamic and kinetic equilibrium. If we follow the time-dependent progression of a growing and splitting assembly, we observe quasi-stationary states (QSSs) when for a few generations the composition of the assembly remains basically the same, showing that molecular compositions can be preserved from one generation to another, in other words, there is a kind of heredity. The relatively stable compositions that can be maintained for generations are called compotypes. Due to the stochastic growth and division abrupt changes from one QSS to another appear that correspond to mutation of compotypes. Therefore, the information transmitted across generations is the molecular composition of the entity, a phenomenon called compositional inheritance.

The most important feature of GARD is that it demonstrates compositional inheritance (Segré et al., 2000, 2001b), but it must be emphasized that inheritance was analyzed by following the fate