Journal of Theoretical Biology 364 (2015) 249–259

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

The origin and spread of a cooperative replicase in a prebiotic chemical system

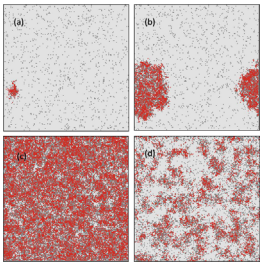
Julie A. Shay [1], Christopher Huynh, Paul G. Higgs *

Department of Physics and Astronomy, McMaster University, Hamilton, Ontario, Canada L8S 4M1

HIGHLIGHTS

- We model the origin of a trans-acting replicase in a prebiotic chemical system.
- We study alternating plus/minus strand replication.
- Pre-existing random sequences act as parasitic templates.
- A replicating state can sometimes spread from an isolated initial replicase.
- A mean field theory explains the differences between spatial and well-mixed models.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 9 June 2014
Received in revised form
19 August 2014
Accepted 12 September 2014
Available online 22 September 2014

Keywords:
Origin of Life
RNA World
Autocatalysis
Prebiotic chemistry
Template-directed polymerization

ABSTRACT

The origin of life requires the emergence of a system of autocatalytic polymers such as RNA. We consider a trans-acting replicase that catalyses replication of a template (either a copy of itself or another sequence). Our model includes alternating plus/minus strand replication where only the plus strand is a catalyst. Prebiotic chemistry generates random sequences and allows for non-catalysed, template-directed synthesis of new strands. These chemical reactions are insufficient to sustain replication, but they provide a background in which the first replicase can arise. In the well-mixed case, the minimum value of the catalytic rate parameter k for which a stable replicating state survives scales as $1/f$, where f is the fraction of random sequences that are catalysts. When catalysts are rare ($f \to 0$), the replicating state is not stable in for any finite k because the replicases are overrun by parasitic templates already present in the prebiotic system, and by additional parasites created by mutation of the catalyst. In contrast, in 2d spatial simulations, the replicating state is stable for moderate k with appropriate values of the local diffusion constant. We calculate the probability of spread of the replicating state from a single isolated catalyst. This occurs in a parameter range that is narrower than that in which existing replicators are stable. The 2d model uses 'Two's Company' rules, where two molecules on a site may replicate, but crowding occurs when three molecules are on one site. A mean-field theory is presented which predicts the most important results of the spatial model. Our results emphasize that the origin of replication is a spatially-localized stochastic transition between a 'dead' state controlled by prebiotic chemistry and a 'living' state controlled by autocatalytic replication.

1. Introduction

A key step in the origin of life is the formation of an autocatalytic system of replicating molecules. It seems highly likely that RNA

* Corresponding author. Tel.: +905 525 9140x26870.
E-mail address: higgsp@mcmaster.ca (P.G. Higgs).
[1] Present address: Molecular Biology and Biochemistry Department, Simon Fraser University, Burnaby, BC, Canada V5A 1S6.

http://dx.doi.org/10.1016/j.jtbi.2014.09.019
0022-5193/

preceded DNA and proteins in living systems, and that it had both genetic and catalytic roles in the early stages of life (Gilbert, 1986; Bartel and Unrau, 1999; Joyce, 2002). The first self-replicating molecules may have been made of RNA, or a similar nucleic acid analogue that forms complementary sequence pairs (Hud et al., 2013). The RNA World is still the foremost hypothesis in our understanding of the origin of life (Bernhardt, 2012).

In order for the RNA world to have existed, there must have been a ribozyme capable of catalysing its own replication. A series of polymerase ribozymes of increasing performance has been created *in vitro* (Johnston et al., 2001; McGinness et al., 2002; Zaher and Unrau, 2007; Wochner et al., 2011) and autocatalytic replication of RNA systems based on ligases (Lincoln and Joyce, 2009) and recombinases (Hayden et al., 2008) has also been demonstrated. Although none of these examples yet qualifies as a replicating system that could sustain itself using only chemical resources that might be found in the prebiotic world, this research gives a strong suggestion that such a ribozyme could exist. In this paper, we will discuss the origin and stability of replicating molecules. We have RNA systems in mind, but the models could also apply to other polymers that may have developed a replicating system preceding the RNA world (Hud et al., 2013; Orgel, 2004), if these had the ability to be both catalysts and templates with complementary pairing.

There has been a lot of previous work on theoretical models of replication, and we wish to make the similarities and differences clear between the system we study in this paper and previous models. The simplest kind of replication is where one molecule makes a second copy of itself, which we can denote as $X \rightarrow 2X$. If the possibility of mutation is allowed for so that the second molecule may not be an identical copy of the first, then a population of related sequences emerges. This is the standard quasispecies theory for molecular evolution (Eigen et al., 1988). In this picture, each molecule has a fitness (*i.e.* a replication rate) that is a property of its own sequence. Experiments on RNA replication with Qβreplicase (Biebricher et al., 1985) provide a good example. The fitness in this case is the rate at which a sequence is replicated by the replicase enzyme, which is supplied in the experiment and is not coded by the sequence.

In the RNA world, however, we are interested in replication that is catalysed by another RNA sequence, which we can denote as $2X \rightarrow 3X$. Here one copy of the sequence acts as a replicase and one acts as a template, and together they make a third copy. In this case it is important to distinguish between the ability of a sequence to be a template and its ability to be a replicase. The efficiency of a sequence as a template may depend on which other sequences are acting as replicases, hence the 'fitness' of a sequence is not simply a property of its own sequence. The simplest case of this type is a model in which there is a replicase (X) that copies all other sequences with equal ability. In addition to copying itself ($2X \rightarrow 3X$), it may copy another molecule (Y) to make a second copy of the other molecule ($X+Y \rightarrow X+2Y$). The X molecules are co-operators because they spend time replicating the Y sequences at no benefit to themselves, whereas the Y sequences are parasites because they cannot replicate themselves. In a well-mixed system, the replicase is only marginally stable if replication is completely accurate, whereas if errors in replication of X produce additional Ys, then the parasites overrun the system and the replicase does not survive (McCaskill et al., 2001; Brogioli, 2010; Takeuchi and Hogeweg, 2012). The outcome is different if spatial structure is added into these models, *e.g.* by considering molecules diffusing on a two dimensional surface and interacting with their neighbours. A large number of different theoretical models with two-dimensional diffusion show that clustering of co-operating replicases allows the replicases to survive under conditions in which they would be destroyed by parasites in a well-mixed system (McCaskill et al., 2001; Brogioli, 2010; Takeuchi and Hogeweg, 2012, 2007; Boerlijst and Hogeweg, 1991, 1995; Chacón and Nuño, 1995; Cronhjort and Blomberg, 1997; Szabo et al., 2002; Hogeweg and Takeuchi, 2002; Fuchslin et al., 2004; Sardanyés and Solé, 2007; Sardanyés, 2008).

Despite this large literature, most of these papers do not directly address what we consider to be the most relevant situation for the RNA World – that is trans-acting replicases that use alternating plus/minus strand replication, and where only the plus strand is a catalyst. Here, we will build on the simplest case of X and Y molecules described above that was studied in McCaskill et al. (2001). In that work, complementary sequences were ignored for simplicity. In our work, the plus strand X is a replicase, and the complementary minus strand W has no catalytic function. Both X and W can be templates. The replication processes are $2X \rightarrow 2X+W$, and $X+W \rightarrow 2X+W$. Replication of parasitic Y sequences by X is also included in our model. As the catalyst is assumed to replicate all sequences equally well, Y is any other sequence that is not X or W.

One kind of replicating system that has been extensively studied is the hypercycle (Boerlijst and Hogeweg, 1991, 1995; Chacón and Nuño, 1995; Sardanyés and Solé, 2007; Sardanyés, 2008). In hypercycle models, there are multiple catalysts, and each one specifically catalyses replication of the next one in the cycle. Reaction steps are of the form $X_{i-1}+X_i \rightarrow X_{i-1}+2X_i$, meaning that X_{i-1} is the catalyst for X_i and that another copy of X_i is formed (not the complement of X_i). A two component hypercycle would be $X_1+X_2 \rightarrow X_1+2X_2$, and $X_2+X_1 \rightarrow X_2+2X_1$. We emphasize that the system with complementary X and W strands is *not* a two-component hypercycle. Previous work on hypercycles is relevant to the extent that it shows the importance of spatial structure as a means of resisting parasites, but the RNA World model we study here is mathematically distinct from hypercycles and needs to be investigated separately.

Another point that should be noted about most previous spatial models of replication is that these focus almost exclusively on the stability of existing replicases and not on the origin of the replicating system in the first place. Our own work has addressed the way that autocatalytic replicase molecules could emerge in a system of prebiotic chemistry that is capable of synthesizing random sequences (Wu and Higgs, 2009, 2011, 2012; Higgs and Wu, 2012). The essential feature of our models is a replicating molecule that can be formed either by spontaneous reactions (prebiotic chemistry) or by an autocatalytic process (biology). Several different models of this type were studied using differential equations that describe the concentrations of the different types of molecules in these chemical reaction systems. These models share the important feature of having two types of stable solution that we refer to as dead and living. In the dead state there is a very low (or sometimes zero) equilibrium concentration of catalysts that is produced by spontaneous chemistry. In the living state, there is a high equilibrium concentration of catalysts that is sustained by self-replication. The chemical reaction terms that produce the first few catalysts are necessary for the origin of the living state but they are not necessary for its continued survival because the autocatalytic reactions are much faster than the non-living chemical reactions once a high concentration of catalysts is established.

An important result from our previous papers (Wu and Higgs, 2009, 2011, 2012; Higgs and Wu, 2012) is that the origin of replication is a stochastic process. In a deterministic reaction system with no concentration fluctuations, a system beginning in the dead state will remain dead forever. In a finite volume with finite numbers of molecules, stochastic concentration fluctuations can cause the system to jump into the autocatalytic replicating state (Wu and Higgs, 2009). In our view, the existence of autocatalytic