Generic Darwinian selection in catalytic protocell assemblies

Andreea Munteanu[1,*], Camille Stephan-Otto Attolini[2], Steen Rasmussen[3,4], Hans Ziock[3] and Ricard V. Solé[1]

[1]*ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Dr Aiguader 80, 08003 Barcelona, Spain*
[2]*Universität Wien, Währingerstraße 17, 1090 Wien, Austria*
[3]*SOS, EES6 MS 462 Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[4]*Santa Fe Institute, Santa Fe, NM 87501, USA*

To satisfy the minimal requirements for life, an information carrying molecular structure must be able to convert resources into building blocks and also be able to adapt to or modify its environment to enhance its own proliferation. Furthermore, new copies of itself must have variable fitness such that evolution is possible. In practical terms, a minimal protocell should be characterized by a strong coupling between its metabolism and genetic subsystem, which is made possible by the container. There is still no general agreement on how such a complex system might have been naturally selected for in a prebiotic environment. However, the historical details are not important for our investigations as they are related to assembling and evolution of protocells in the laboratory. Here, we study three different minimal protocell models of increasing complexity, all of them incorporating the coupling between a 'genetic template', a container and, eventually, a toy metabolism. We show that for any local growth law associated with template self-replication, the overall temporal evolution of all protocell's components follows an exponential growth (efficient or uninhibited autocatalysis). Thus, such a system attains exponential growth through coordinated catalytic growth of its component subsystems, independent of the replication efficiency of the involved subsystems. As exponential growth implies the survival of the fittest in a competitive environment, these results suggest that protocell assemblies could be efficient vehicles in terms of evolving through Darwinian selection.

Keywords: protocell; replicator dynamics; catalytic coupling; prebiotic evolution

1. INTRODUCTION

Cells are the basic structural and functional units of all known life, performing the vital functions of an organism and containing the hereditary information necessary for self-regulation and self-replication. The complexity of modern cells and life itself is a consequence of open-ended evolution whose beginnings are associated with the origin of life on earth. Even though many of the main issues associated with the emergence of life remain to be established, significant progress has been made since Oparin's original vision of the origin of life (Oparin 1957) and Gánti's model of a minimal cell (Gánti 1975).

These advances have steadily pushed the experimental implementation as well as the understanding of protocells ever closer to reality (Kaneko & Yomo 2000; Segré *et al*. 2001; Szostak *et al*. 2001; Monnard *et al*. 2002; Stadler & Stadler 2003; Takakura *et al*. 2003; Paul & Joyce 2004; Rasmussen *et al*. 2004; Luisi *et al*. 2006). For a protocell to function properly, it must contain both genes and a metabolism, which are integrated by a container. The interdependence

and cross-regulation of these subsystems are major issues. Regardless of which of the protocellular subsystems a work centres on—templates, energetics or container—the concept of self-replication is central, be it an experimental scenario or a theoretical model.

One important problem is the overall kinetic behaviour of template replication. Information in biological systems is stored in, and is propagated by, template molecules capable of making copies of themselves. Owing to the product inhibition by the new complementary templates in a naked gene replication process, the kinetics have been shown to be parabolic (Bag & von Kiedrowski 1996), i.e. subexponential, and prevent Darwinian selection. Since exponential growth is known to be a key feature of Darwinian selection, where the fittest variants win over the less-fit ones, the lack of exponential growth for reproducers would be an obstacle to their success. In spite of our intuition, which might suggest that selection at the cellular level should become slower than parabolic when based on parabolic template dynamics, here we show quite the opposite: the selection process in catalytically coupled assemblies of protocells is Darwinian, with coordinated exponential growth of all involved components and thus survival of the fittest, independent of the template growth law. The present work investigates a class of minimal catalytically coupled protocell models and is centred on the

* Author for correspondence (andreea.munteanu@upf.edu).

relation between the internal dynamics and the global population-level dynamics of protocells.

2. BASIC GROWTH DYNAMICS

From simple models of prebiotic evolution, a few basic growth laws resulted as being fundamental for replicator's dynamics. These, in turn, have remarkable implications for the selection among competing processes (Eigen & Schuster 1979; Szathmáry 1991). The primary replication process, the fission of the progenitor, gives rise to exponential growth in the absence of ecological constraints (Malthusian growth). Here, the growth rate of the population is proportional to the population itself, i.e. $\dot{x} = kx$, where '$\cdot$' implies a time derivative. When two exponentially growing populations compete, the one with the greater growth rate completely excludes its competitor, a case referred to as 'the survival of the fittest'. Additionally, two other simple non-Malthusian processes are often considered in the literature: a faster and a slower growth than the exponential one (table 1). While the former, generally referred to as hyperbolic growth ($\dot{x} = kx^p$, $p > 1$), is associated with models of mutualistic replicators (hypercycles; Eigen 1971), the latter (parabolic growth: $\dot{x} = kx^p$, $p < 1$) is typical for experimental template-directed replication ($p = 0.5$; von Kiedrowski 1986). Note that constant growth is also possible in this framework ($p = 0$).

One can note that for two coexistent parabolically growing populations[1], 'the survival of everybody' is guaranteed, while for two hyperbolically[2] growing populations, the one with the highest initial concentration times the growth rate (i.e. $x_0 k$) will outgrow the other (Eigen & Schuster 1979; Szathmáry & Maynard-Smith 1997). While referred to as 'the survival of the common' case, it is not guaranteed that the most common hyperbolically growing population will always invade when rare. Which one wins is determined by the ratio between initial populations (initial conditions), on one hand, and the fitness constants (reaction constants), on the other hand (consider the case $p = 2$ in table 1). Although Darwinian evolution, i.e. survival of the fittest, is possible for the hyperbolic case, the clearest example of Darwinian evolution is provided by exponential growth.

3. TEMPLATE-DIRECTED REPLICATION

Among all systems with auto-catalytic synthesis of their constituent molecules, the most relevant to the origin of life is the one based on non-enzymatic template-directed replication (Orgel 1992). In the last two decades, various experimental works have been dedicated to the in-depth study of such simple systems with emphasis on artificial 'self-replication' (Robertson *et al.* 2000). Several research groups were able to experimentally demonstrate a sigmoidal time evolution of template concentration (von Kiedrowski *et al.* 1991; Lee *et al.* 1996). Such behaviour is a consequence of the exponential growth of the product (i.e. template) in the presence of a limited number of available building blocks (limited resources). Von Kiedrowski and co-workers were the first to obtain self-complementary artificial replication (von Kiedrowski 1986) and showed

Table 1. Growth laws.

growth law general solution	$\dot{x} = kx^p$, $p > 0$
	$x(t) = \begin{cases} \left[x_0^{1-p} + (1-p)kt\right]^{1/1-p} & \text{for} \quad p \neq 1 \\ x(t) = x_0 e^{kt} & \text{for} \quad p = 1 \end{cases}$
particular cases	$p = 0$: constant
	$p = 1/2$: pure parabolic
	$p = 1$: exponential
	$p = 2$: pure hyperbolic

a growth order $p = 1/2$ (parabolic growth) instead of $p = 1$ (exponential growth) as expected from a true auto-catalytic replication. Parabolic growth has been proved to be the result of product inhibition that decreases the efficiency of the auto-catalytic cycle (Bag & von Kiedrowski 1996).

As described in §2, the auto-catalytic growth order is of great importance in establishing the long-term evolution of the system through selection processes. Interestingly enough, Rocheleau *et al.* (2007) proved that in spite of an internal parabolic growth of template concentration in a simple model of template–container coupling, the total template concentration considered over the entire population of protocells turns out to be exponential. More surprisingly, the catalytically coupled protocell containers follow an exponential growth with the same exponent as the templates' growth. This result is significant as a proof of regulated template–container interdependence, ensuring a correlated growth of the system's parts. This work thus demonstrates that a stoichiometric coupling between the protocell's subsystems is not compulsory, as suggested by others (Gánti 2003) and that cross-regulation of the subsystems occurs naturally as a consequence of catalysis. In the present work, we shall demonstrate that this result holds for any type of local growth dynamics of templates in the model of Rocheleau *et al.* (2007), as well as in extended and more realistic versions that include metabolic reaction steps.

4. TEMPLATE–CONTAINER COUPLING

In the work by Rocheleau *et al.* (2007), the template–container system was effectively defined by the chemical reactions

$$\alpha O + T_s \xrightarrow{k_T} T_d, \tag{4.1a}$$

$$T_d \overset{K_t}{\leftrightarrow} 2T_s, \tag{4.1b}$$

$$pL + T_d \xrightarrow{k_L} T_d + L, \tag{4.1c}$$

where T_s, T_d and L represent the single-strand template, the double-strand template and the lipid molecules, respectively, while O and pL are, respectively, the single-stranded template and lipid precursors whose concentrations are kept constant, with α denoting the necessary stoichiometry for the template-directed replication. The graphical representation of the chemical reactions appears in figure 1. The hybridization and the dehybridization are considered in equilibrium (reaction (4.1b)) as long as the template replication (reaction (4.1a)) is a
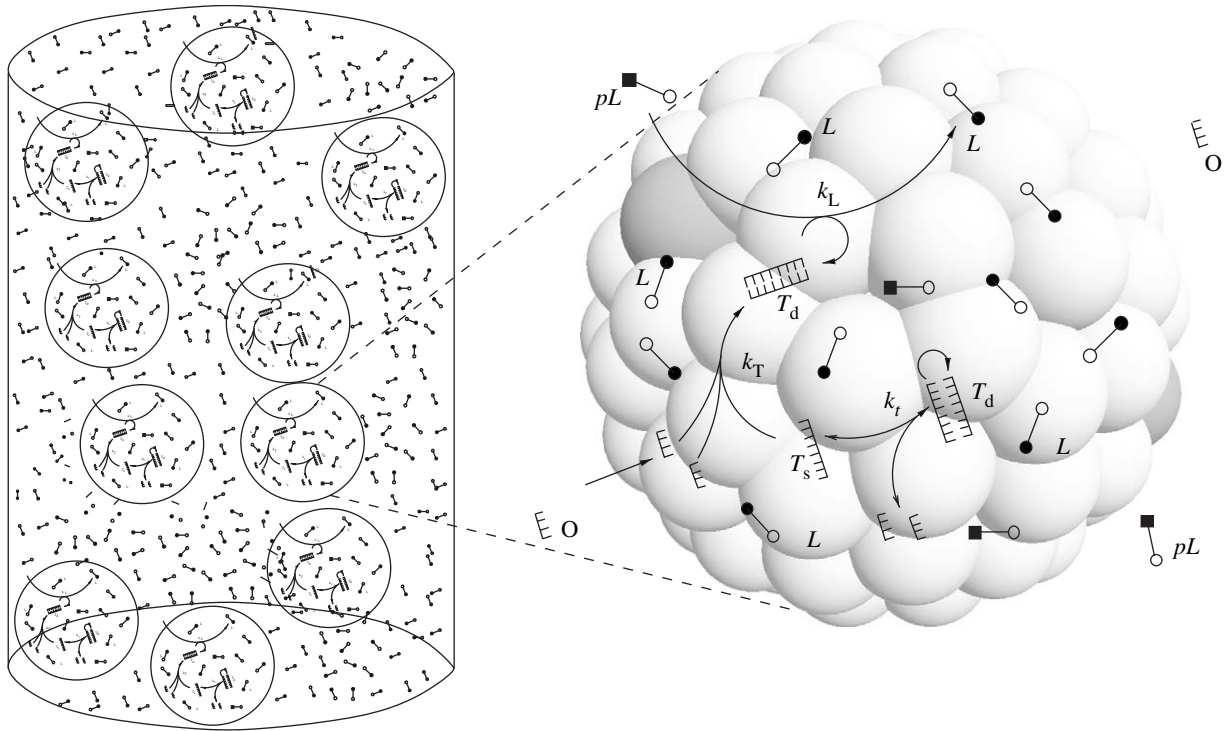
Figure 1. A simple model of protocell replication involving template-catalysed formation of lipid molecules. This is the simplest scenario of coupled template–container organization. The circular features representing the individual protocells consist of a micellar container built of lipid molecules (L) among which are embedded the templates (T: s = single stranded, d = double stranded), the lipid precursors (pL) and the resource template oligomers (O). The 'k's are the rate constants.

significantly slower process in comparison. The former condition leads to the equilibrium constant $K_t = [T_s]^2/[T_d]$, where the brackets denote concentration (molarity). As expected from experimental work, when dehybridization is much slower than hybridization (i.e. product inhibition), one has $K_t \ll 1$ since $[T_d] \gg [T_s]$.

Rocheleau *et al.* (2007) define the total single-strand concentration as $[T_1] \equiv 2[T_d] + [T_s]$ and consider that the protocells grow by the intake of nutrients (i.e. precursors) and divide when a predefined critical size is reached. The existence of a critical size for micellar structures is supported by both experiments and theory (Bachman *et al.* 1992; Mavelli 2004), even if the growth–division process still lacks a complete experimental and theoretical proof. Additionally, the aggregates are envisioned as composed mainly of lipids, with the other molecules existing at very low concentrations diluted in the local lipid solution.

(a) *Constant volume*

By applying the standard chemical kinetics to the reactions (4.1a)–(4.1c), one finds that the local total template and lipid concentrations evolve according to

$$[\dot{T}_1] = C_T[T_1]^p \quad \text{with} \quad p = 1/2, \qquad (4.2a)$$
$$[\dot{L}] = C_L[T_1], \qquad (4.2b)$$

where $C_T \equiv k_t f([O])\sqrt{K_t/2}$ for $p = 1/2$, and $C_L \equiv k_L[pL]/2$ are constants[3] and '·' implies a time derivative (see electronic supplementary material A). The term $f([O])$ is a function depending on oligomer stoichiometry (the number of oligomers) needed for each template replication. As we consider that the oligomers are present in abundance and thus the

oligomer concentration can be approximated to constant, this function term is also a constant. Local refers to the concentration within the container itself. Considering a general case of equation (4.2a), the local dynamics of the templates is parabolic if $p < 1$, exponential if $p = 1$ and hyperbolic if $p > 1$; while, by equation (4.2b), the lipid local concentration rate remains proportional to the template concentration. We shall demonstrate in this section that independent of the value of p, the global dynamics will be exponential in form although the doubling time or the cycle period will depend on the value of p.

Equations (4.2a) and (4.2b) implicitly assume that the concentration dilution due to the increase of an individual protocell's volume in the growth–division process is negligible. In other words, the protocell's volume (V_1) can be approximated as a constant and equal to a time-averaged protocell's volume (V_A, a true constant). In the first approximation, the concentrations are calculated with respect to the latter volume (i.e. $[X] = N_X/V_A$, where N_X is the number of molecules of type X).

(b) *Growing–dividing protocells*

In a second approximation, Rocheleau *et al.* (2007) consider that the local lipid aggregate volume V_1 grows only from the addition of lipid molecules to the aggregate, i.e. the contribution of templates, precursors and fluid to the aggregate volume was neglected. Note that in the case of a micelle, as considered here, both the volume and the surface area (S) of the protocell are proportional to the number of lipids ($V_1 \propto S \propto N_L$). That is, as the micelles have no hollow interior, the protocell's volume equals the number of lipids multiplying the volume of a lipid. In this case, the

surface equals the number of lipids multiplying πR^2, with R being the radius of a lipid, and thus the surface too is proportional with the number of lipids. In the case of a vesicle, the hollow spherical form introduces a nonlinear dependence between volume and surface $(V_l \propto S^{3/2}; S \propto N_L)$[4].

In the Rocheleau approach, once the dilution effect is taken into account, the concentration of the metabolites must be referred to the time-dependent volume V_l instead of V_A and thus equation (4.2a) must be replaced by

$$\frac{d[T_l]}{dt} = \frac{d[T_l]}{dt}\bigg|_{V_l} - \frac{[T_l]}{V_l}\frac{dV_l}{dt}\bigg|_{N_T} = C_T[T_l]^p - \frac{[T_l]}{V_l}\frac{dV_l}{dt},$$

$$(4.3)$$

where the first term on the r.h.s. is calculated at constant volume $(V_l = V_A)$, while the second at a constant number of template molecules (see electronic supplementary material B). As the lipid aggregates have a characteristic scale, there exists an average number m_0 of lipids per aggregate volume V_A. Theoretically and experimentally, the protocells are expected to grow and become unstable when reaching a critical size, with the instability being resolved by the subsequent division into two daughters.

(c) *Global concentrations*

In the context of growing–dividing protocells in a tank reactor, the global concentration of metabolites denotes their concentration calculated using the volume of the tank containing the aggregates (figure 1). The global concentration of templates, $[T_g]$, is thus obtained in Rocheleau *et al.* (2007) as being

$$[T_g] = [T_l][A]V_l \approx [T_l][A]V_A,\qquad (4.4)$$

where $[A]$ is the concentration of aggregates or protocells (number of aggregates per total tank volume), with the local concentration still referring to the intra-aggregate metabolite concentration.

The growth rate of the aggregate concentration, $d[A]/dt$, depends on both the aggregate concentration, $[A]$, and the growth of the local volume of the aggregates, dV_l/dt. The former is a consequence of the division process, since the greater the number of aggregates that exist, the more will be produced in a finite time-interval, while the latter term is an indicator of the doubling time as it illustrates how fast a single aggregate (and implicitly the number of aggregates) grows through the addition of lipids (see also electronic supplementary material B)

$$\frac{d[A]}{dt} = [A]\frac{d(N_L/m_0)}{dt} = [A]\frac{1}{V_A}\frac{dV_l}{dt}.\qquad (4.5)$$

Taking the derivative of equation (4.4) and using equations (4.4) and (4.5) (see the electronic supplementary material) and $V_l \approx V_A$, the global dynamics are given by

$$\frac{d[T_g]}{dt} = \gamma_T[T_g]\left(\frac{[T_g]}{[A]}\right)^{p-1},\qquad (4.6)$$

$$\frac{d[A]}{dt} = \gamma_A[T_g],\qquad (4.7)$$

where $\gamma_T \equiv C_T/V_A^{p-1}$ and $\gamma_A \equiv C_L/m_0$. See electronic supplementary material B for the detailed derivation of these equations.

Dividing equation (4.6) by equation (4.7), one finds that

$$\frac{d[T_g]}{d[A]} = \frac{\gamma_T}{\gamma_A}\left(\frac{[T_g]}{[A]}\right)^{p-1},\qquad (4.8)$$

which when solved leads to

$$[A] = \left[\frac{\gamma_A}{\gamma_T}[T_g]^{2-p} + C_0\right]^{1/(2-p)}; \quad p \neq 2,\qquad (4.9a)$$

$$[A] = C_1[T_g]^{\gamma_A/\gamma_T}; \quad p = 2.\qquad (4.9b)$$

The constants C_0 and C_1 are defined as $C_0 \equiv [A]_0^{2-p} - (\gamma_A/\gamma_T)[T_g]_0^{2-p}$ and $C_1 \equiv [A]_0/[T_g]_0^{(\gamma_A/\gamma_T)}$, where $[A]_0$ and $[T_g]_0$ denote the aggregate and the template concentrations at $t=0$, respectively. From its definition, $\gamma_A > 0$ and, as $[T_g]$ must be positive, thus $d[A]/dt$ is always positive (see equation (4.7)). Therefore, the concentration $[A]$ increases in time, independent of the sign of constant C_0. However, the general condition for $C_0 > 0$ is

$$(2-p)\ln\left(V_A\frac{[T_g]_0}{[A]_0}\right) < \ln\alpha \quad \text{with} \quad \alpha \equiv \frac{C_T m_0}{C_L V_A}.\qquad (4.10)$$

Equation (4.3) can now be rewritten as

$$\frac{d[T_g]}{dt} = \gamma_T\left(\frac{\gamma_A}{\gamma_T}\right)^{(1-p)/(2-p)}[T_g]\left[1 + \frac{\gamma_T}{\gamma_A}\frac{C_0}{[T_g]^{2-p}}\right]^{(1-p)/(2-p)};$$

$$p < 2,$$

$$(4.11a)$$

$$\frac{d[T_g]}{dt} = c[T_g]^{2-(\gamma_A/\gamma_T)}; \quad p = 2,\qquad (4.11b)$$

$$\frac{d[T_g]}{dt} = \gamma_T C_0^{(1-p)/(2-p)}[T_g]^p\left[1 + \frac{\gamma_A}{\gamma_T}\frac{1}{C_0[T_g]^{p-2}}\right]^{(1-p)/(2-p)};$$

$$p > 2,$$

$$(4.11c)$$

where $c \equiv \gamma_T/C_1$. The complex form of these equations does not allow a direct integration and thus requires the consideration of an approximated solution.

Let us remember that for any exponent β and any real number x satisfying $|x| < 1$, one can use the approximation

$$(1 + x)^\beta = 1 + \beta x + \frac{\beta(\beta-1)}{2!}x^2 + \cdots \quad \approx 1 + \beta x,$$

$$(4.12)$$

with the last approximation being valid for $|x| \ll 1$.

As $[T_g]$ increases with time, one can consider that after a sufficiently long time, the second term in the brackets of equations (4.11a) and (4.11c) reaches values much smaller than unity and thus the expansion (4.12) can be applied, leading to

$$\frac{d[T_g]}{dt} \approx a[T_g] + b[T_g]^{p-1}; \quad p < 2,\qquad (4.13a)$$

$$\frac{d[T_g]}{dt} \approx g[T_g]^2 + h[T_g]^p; \quad p > 2,\qquad (4.13b)$$

where $a \equiv \gamma_T(\gamma_A/\gamma_T)^{(1-p)/(2-p)}$, $b \equiv \gamma_T C_0((1-p)/(2-p))(\gamma_A/\gamma_T)^{(-1/(2-p))}$, $g \equiv \gamma_A C_0^{(-1/(2-p))}(1-p)/(2-p)$ and $h \equiv \gamma_T C_0^{((1-p)/(2-p))}$. We note that if $C_0 > 0$ for any p, then the constants a, g and h are always positive, with the exception of b, which is negative for $p \in (1, 2)$.

Before proceeding, we remark that there is no known real replication process that achieves a growth order $p > 2$. The most efficient replication process known so far involves the hypercycle's dynamics and thus $p = 2$. For this reason, we shall concentrate in the following on the cases $p < 2$.

One can recognize equation (4.13a) as being a Bernoulli equation, whose general solution is known and thus the global concentration of templates for $p \leq 2$ is found to be

$$[T_\mathrm{g}]_t = \frac{\{\exp[(2-p)at + D_0] - b\}^{1/(2-p)}}{a^{1/(2-p)}}; \quad p < 2, \tag{4.14a}$$

$$[T_\mathrm{g}]_t = (D_1 t + D_2)^{1/((\gamma_\mathrm{A}/\gamma_\mathrm{T})-1)}; \quad \gamma_\mathrm{A} \neq \gamma_\mathrm{T}; \quad p = 2, \tag{4.14b}$$

$$[T_\mathrm{g}]_t = [T_\mathrm{g}](0)\mathrm{e}^{ct}; \quad \gamma_\mathrm{A} = \gamma_\mathrm{T}; \quad p = 2, \tag{4.14c}$$

where $D_0 = \ln[b + a[T_\mathrm{g}]_0^{2-p}]$, $D_1 \equiv c((\gamma_\mathrm{A}/\gamma_\mathrm{T}) - 1)$ and $D_2 \equiv [T_\mathrm{g}]_0^{(\gamma_\mathrm{A}/\gamma_\mathrm{T})-1}$. It can be seen that the case $p = 1/2$ coincides with eqn (30) from Rocheleau *et al.* (2007). Also, similar to the reasoning from Rocheleau *et al.* (2007), for the case $p < 2$, the global template concentration evolves at large times as $[T_\mathrm{g}]_t \approx \exp(at)$ (see also their eqn (21)), with the subscript t denoting the time dependence of the global concentrations. Thus, for $p < 2$ and for $p = 2$, $\gamma_\mathrm{T} = \gamma_\mathrm{A}$, the functional forms of $[A]_t$ and $[T_\mathrm{g}]_t$ after sufficiently long time differ by a constant (equations (4.9a) and (4.9b)), implying that, to a leading order, the coupled template and aggregate growths are both exponential with the same exponent.

More exactly, considering that the constant C_0 becomes negligible when time and concentration get to be large, the ratio of the templates and aggregates concentration (from equations (4.9a) and (4.9b)) behaves as

$$\frac{[T_\mathrm{g}]_t}{[A]_t} \xrightarrow{t \to \infty} \left(\frac{\gamma_\mathrm{T}}{\gamma_\mathrm{A}}\right)^{1/(2-p)}, \tag{4.15}$$

namely the template to aggregate ratio becomes a constant. Next, we remark that as $[T_1] \equiv 2[T_\mathrm{d}] + [T_\mathrm{s}]$ with $[T_\mathrm{d}] \gg [T_\mathrm{s}]$, a viable aggregate must have at least one single template or, in a looser version, one double template, otherwise its replication is impossible. Hence, an implicit requirement for a viable system is $[T_\mathrm{g}]/[A] \geq 1$ or

$$\gamma_\mathrm{A} \leq \gamma_\mathrm{T}, \tag{4.16}$$

for any system characterized by $p < 2$. An example of a simulation that follows equations (4.6) and (4.7) is given in figure 2. We have considered the following values: $m_0 = 1000$, $k_\mathrm{L} = 25\ \mathrm{Ms}^{-1}$, $k_\mathrm{T} = 10\ \mathrm{Ms}^{-1}$, $[pL] = 10^{-3}\ \mathrm{M}$, $V_\mathrm{A} = 5 \times 10^{-19}\ \mathrm{l}$, $K_t V_\mathrm{A} = 1 \times 10^{-3}\ \mathrm{mol}$, $f(O) = \alpha[O] = 2 \times 10^{-4}\ \mathrm{M}$, with $\alpha = 2$ being the number of oligomers necessary in the replication of T_s. The initial conditions are $[T_\mathrm{g}]_0 = 10^{-6}\ \mathrm{M}$ and $[A]_0 = 10^{-8}\ \mathrm{M}$.

However, the case $p = 2$, $\gamma_\mathrm{T} \neq \gamma_\mathrm{A}$ does not present coordinated growth, but instead the aggregate concentration grows faster than the template concentration. For this case, using equations (4.9a) and (4.14b), one
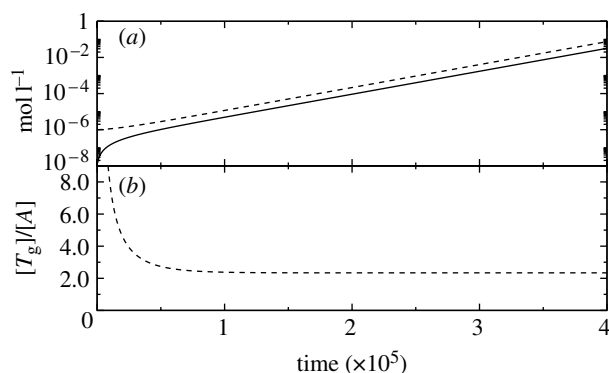

Figure 2. A simulation of equations (4.6) and (4.7) for the parabolic case $p = 1/2$. (*a*) $[A]$, solid line; and $[T_\mathrm{g}]$, dashed line. (*b*) The ratio $[T_\mathrm{g}]/[A]$ satisfies equation (4.15). See text for details.

obtains

$$\frac{[T_\mathrm{g}]_t}{[A]_t} = \frac{1}{C_1}\left[(D_1 t + D_2)^{1/(\gamma_\mathrm{A}/\gamma_\mathrm{T}-1)}\right]^{1-\gamma_\mathrm{A}/\gamma_\mathrm{T}}, \tag{4.17}$$

$$\frac{[T_\mathrm{g}]_t}{[A]_t} = \frac{1}{C_1(D_1 t + D_2)} \xrightarrow{t \to \infty} 0, \tag{4.18}$$

confirming a faster growth of the aggregate concentration with respect to the template concentration. The physical interpretation of this result is that, eventually, there would exist many aggregates that contain no templates at all.

Finally, for the case $p > 2$, one can note that at large global template concentrations, equation (4.13b) can be approximated further as

$$\frac{\mathrm{d}[T_\mathrm{g}]}{\mathrm{d}t} \approx h[T_\mathrm{g}]^p, \tag{4.19}$$

and thus the global template concentration follows the local intra-aggregate growth law, with order p. As such, the global growth law is no longer exponential. Nonetheless, as previously discussed, after a sufficiently long time, equation (4.9a) reveals that $[A]_t$ and $[T_\mathrm{g}]_t$ differ by a constant. Thus, for the cases $p > 2$, the resultant coordinated growth of the population of protocells is characterized by the same exponent as the local intra-aggregate template growth.

Once again, we stress the fact that no known real growth process has an order $p > 2$ and, in fact, all theoretical and experimental replication processes present growth orders inferior to the pure hyperbolic case ($p = 2$). We consider as our most important result the derivation that for the real world ($p < 2$), the coupling of the template with the container in a protocellular ensemble results in a coordinated global exponential growth at large times.

5. THREE-ELEMENT DYNAMICS

A higher level of chemical complexity can be introduced by considering additional reactions for the template replication. This seems a reasonable assumption since experimental template replication requires the activation of the precursors subsequently employed in the replication process. In this sense, we shall consider the introduction of oligomer precursors, pO into the
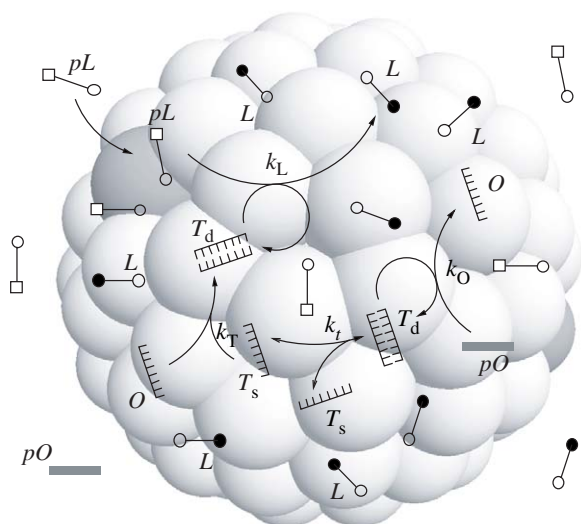
Figure 3. Protocell reaction network for the three-element scenario. In this case, the container is coupled to the template through an additional reaction structure where precursors of templates need to be transformed, the transformation being catalysed by templates. For key see figure 1.

dynamics of the two-element system presented above. The chemical reactions considered are

$$pO + T_d \xrightarrow{k_O} O + T_d, \qquad (5.1a)$$

$$\alpha O + T_s \xrightarrow{k_T} T_d, \qquad (5.1b)$$

$$T_d \overset{K_t}{\leftrightarrow} 2T_s, \qquad (5.1c)$$

$$pL + T_d \xrightarrow{k_L} L + T_d, \qquad (5.1d)$$

where the notations from Rocheleau *et al.* (2007) have been used: T_s, single-stranded template; T_d, double-stranded template; pO and pL, the oligomer and lipid precursors, respectively. A graphical representation of this protocell's chemistry is shown in figure 3. The steady state is considered, that is the precursors' concentration is taken as constant. In this case, one can note that the double-stranded template catalyses both lipid and oligomer formation. As in the previous section, we use the total local template concentration $[T_l] \equiv 2[T_d] + [T_s] \approx 2[T_d]$ as the template variable and take into account the effect of volume change with time. Considering the kinetics of this system, the concentration of templates and oligomers can be respectively written as

$$[\dot{T}_l] = \alpha k_T [O_l] \sqrt{\frac{K_t [T_l]}{2}} - \frac{d[A]}{dt} \frac{[T_l]}{[A]}, \qquad (5.2a)$$

$$[\dot{O}_l] = k_O [pO] \frac{[T_l]}{2} - k_T [O_l] \sqrt{\frac{K_t [T_l]}{2}} - \frac{d[A]}{dt} \frac{[O_l]}{[A]}. \qquad (5.2b)$$

As in §4, the transition from local to global is based on the expressions $[T_g] = [T_l][A]V_A$ and $[O_g] = [O_l][A]V_A$.
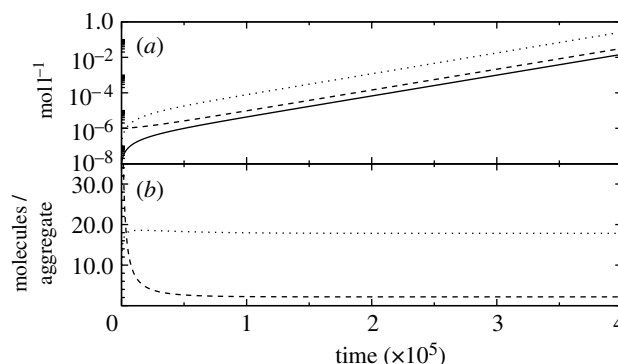
Figure 4. A simulation of equations $(5.3a)$–$(5.3c)$ with $k_O = 5 \text{ Ms}^{-1}$, $k_T = 15 \text{ Ms}^{-1}$ and the rest of parameters having the values employed in figure 2. The values of the final constants are $\gamma_0 = 2.5 \times 10^{-4}$, $\gamma_T = 2.2 \times 10^{-6}$ and $\gamma_A = 1.25 \times 10^{-5}$. (a) $[A]$, solid line; $[T_g]$, dashed line and $[O_g]$, dotted line. (b) The ratio $[T_g]/[A]$, dashed line; and $[O_g]/[A]$, dotted line.

Thus, the final system of coupled equations is

$$[\dot{T}_g] = \gamma_T [O_g] \sqrt{\frac{[T_g]}{[A]}}, \qquad (5.3a)$$

$$[\dot{O}_g] = \gamma_O [T_g] - \gamma_T [O_g] \sqrt{\frac{[T_g]}{[A]}}, \qquad (5.3b)$$

$$[\dot{A}] = \gamma_A [T_g], \qquad (5.3c)$$

with

$$\gamma_O \equiv \frac{k_O [pO]}{2}; \quad \gamma_T \equiv \alpha k_T \sqrt{\frac{K_t}{2V_A}};$$

$$\gamma_A \equiv \frac{k_L [pL]}{2m_0}, \qquad (5.4)$$

where m_0 is again the average number of lipid molecules per aggregate and $K_t = [T_s]^2/[T_d]$ resulting from the equilibrium condition of reaction $(5.1c)$. Figure 4 illustrates an example of a solution of equations $(5.3a)$–$(5.3c)$. The behaviour appears to be representative of the generic dynamics of the entire range of parameters investigated. Below, we shall demonstrate that this case is also characterized by coordinated exponential growth after a sufficiently long time, with the same exponent for the oligomer, template and aggregate concentration.

(a) Linear stability analysis

The nonlinear system from equations $(5.3a)$–$(5.3c)$ has $([O_g], [T_g], [A]) = ([O_g]_0, 0, [A]_0)$ as a unique fixed point that results to be unstable under perturbation, implying that these variables tend to infinity when time tends to infinity. In other words, a protocell whose chemistry is given by equations $(5.1a)$–$(5.1d)$, but containing no template molecules, will not grow. On the contrary, a protocell containing template molecules will grow and replicate, producing the increase in the global metabolites' concentrations. Employing a change of variables, we shall prove below that the global variables follow an exponential growth.

As we start with at least one aggregate and the protocells' death is not contemplated, we conclude that the following variables can be well defined: $x \equiv [T_g]/[A]$ and $y \equiv [O_g]/[A]$. Using these variables, equations $(5.3a)$–$(5.3c)$ reduce to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{[\dot{T}_g][A] - [T_g][\dot{A}]}{[A]^2} = \gamma_T y x^{1/2} - \gamma_A x^2, \qquad (5.5a)$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{[\dot{O}_g][A] - [O_g][\dot{A}]}{[A]^2} = \gamma_O x - \gamma_T y x^{1/2} - \gamma_A x y. \qquad (5.5b)$$

We are interested in studying the dynamics of equations $(5.5a)$ and $(5.5b)$ in order to infer the long-term behaviour of equations $(5.3a)$–$(5.3c)$. In particular, we are interested in whether the ratios x and y stabilize. For this purpose, we note that the former system has two fixed points resulting from the conditions $\dot{x} = 0$ and $\dot{y} = 0$: the trivial one, $(x_1, y_1) = (0, y_0)$, and a second one, (x_2, y_2), which is the solution of the equations

$$\gamma_T y - \gamma_A x^{3/2} = 0; \quad \gamma_O - \gamma_A x - \frac{\gamma_A^2 x^{3/2}}{\gamma_T} = 0. \qquad (5.6)$$

The trivial fixed point illustrates the fact that no evolution is possible in the absence of templates. Thus, we are only interested in the non-trivial case when the initial template concentration is non-zero, i.e. we shall concentrate on the characteristics of the fixed point (x_2, y_2). In order to determine its stability, the Jacobi matrix has to be calculated and evaluated at the fixed point

$$\mathcal{J}_2 \equiv \begin{pmatrix} \dfrac{\partial \dot{x}}{\partial x} & \dfrac{\partial \dot{x}}{\partial y} \\[2mm] \dfrac{\partial \dot{y}}{\partial x} & \dfrac{\partial \dot{y}}{\partial y} \end{pmatrix}_{(x_2, y_2)}$$

$$= \begin{pmatrix} -\dfrac{3}{2}\gamma_A x_2 & \gamma_T x_2^{1/2} \\[2mm] \dfrac{1}{2}\gamma_A x_2 & -\gamma_T x_2^{1/2} - \gamma_A x_2 \end{pmatrix},$$

where equation (5.6) has been employed to simplify the Jacobi matrix, that is we used $y_2 = x_2^{3/2}\gamma_A/\gamma_T$. The eigenvalues λ, which provide the stability information of (x_2, y_2), result from the characteristic equation $\det(\mathcal{J}_2 - \lambda I) = 0$, where I is the identity matrix

$$\lambda^2 + a\lambda + b = 0, \qquad (5.7)$$

$$a \equiv \gamma_T x_2^{1/2} + \frac{5}{2}\gamma_A x_2; \quad b \equiv \gamma_T \gamma_A x_2^{3/2} + \frac{3}{2}\gamma_A^2 x_2^2, \qquad (5.8)$$

$$\Delta \equiv a^2 - 4b = \gamma_T^2 x_2 + \frac{\gamma_A^2 x_2^2}{4} + \gamma_T \gamma_A x_2^{3/2} > 0. \qquad (5.9)$$

It can be seen that since $a, b > 0$ and $\Delta > 0$, the eigenvalues are real and negative (Routh–Hurwitz criterion), leading to the conclusion that the fixed point (x_2, y_2) is stable and implicitly that (x_1, y_1) is unstable.

Returning to equations $(5.3a)$–$(5.3c)$ and to the definition of x and y, one can now see that after a sufficiently long time, the following approximations hold

$$\frac{[T_g]}{[A]} \xrightarrow{t \to \infty} x_2, \qquad (5.10a)$$

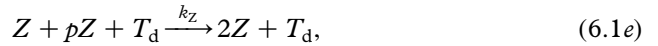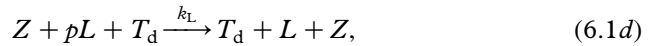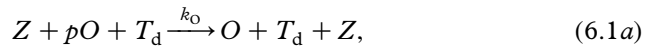$$\frac{[O_g]}{[A]} \xrightarrow{t \to \infty} y_2, \qquad (5.10b)$$

which translate into $[\dot{A}] \approx \gamma_A x_2[A]$ with the use of equation $(5.3c)$ and thus into an exponential increase of the aggregate concentration. Since $[A]$ grows exponentially, and after a sufficiently long time, $[T_g]$, $[O_g]$ and $[A]$ differ by a constant, all the variables grow exponentially with the same exponent. We emphasize once more that this behaviour is satisfied only after a sufficiently long time.

Additionally, one can employ equations (5.6) to show that the second fixed point satisfies the relation: $x_2 + y_2 = \gamma_0/\gamma_A$ (figure 4).

6. FOUR-ELEMENT DYNAMICS

Here, we consider a more complete and realistic protocellular system incorporating an explicit metabolism driven by a light-activated sensitizer molecule (Rasmussen *et al.* 2003). The set of reactions is

$$Z + pO + T_d \xrightarrow{k_O} O + T_d + Z, \qquad (6.1a)$$

$$\alpha O + T_s \xrightarrow{k_T} T_d, \qquad (6.1b)$$

$$T_d \overset{K_t}{\leftrightarrow} 2T_s, \qquad (6.1c)$$

$$Z + pL + T_d \xrightarrow{k_L} T_d + L + Z, \qquad (6.1d)$$

$$Z + pZ + T_d \xrightarrow{k_Z} 2Z + T_d, \qquad (6.1e)$$

where the added components are the sensitizer and its precursor, Z and pZ, respectively. The sensitizer absorbs light energy to drive the conversion of the precursor molecules to their products (figure 5).

Again, the standard kinetic differential equations associated with the global variables of this chemical network are easily recovered:

$$\frac{\mathrm{d}[T_g]}{\mathrm{d}t} = \gamma_T[O_g]\sqrt{\frac{[T_g]}{[A]}}, \qquad (6.2a)$$

$$\frac{\mathrm{d}[O_g]}{\mathrm{d}t} = \gamma_O[Z_g]\frac{[T_g]}{[A]} - \gamma_T[O_g]\sqrt{\frac{[T_g]}{[A]}}, \qquad (6.2b)$$

$$\frac{\mathrm{d}[Z_g]}{\mathrm{d}t} = \gamma_Z[Z_g]\frac{[T_g]}{[A]}, \qquad (6.2c)$$

$$\frac{\mathrm{d}[A]}{\mathrm{d}t} = \gamma_A[Z_g]\frac{[T_g]}{[A]}, \qquad (6.2d)$$

where the constants are defined as

$$\gamma_O \equiv \frac{k_O[pO]}{2V_A}; \quad \gamma_T \equiv \alpha k_T \sqrt{\frac{K_t}{2V_A}},$$

$$\gamma_Z \equiv \frac{k_Z[pZ]}{2V_A}; \quad \gamma_A \equiv \frac{k_L[pL]}{2m_0 V_A}.$$
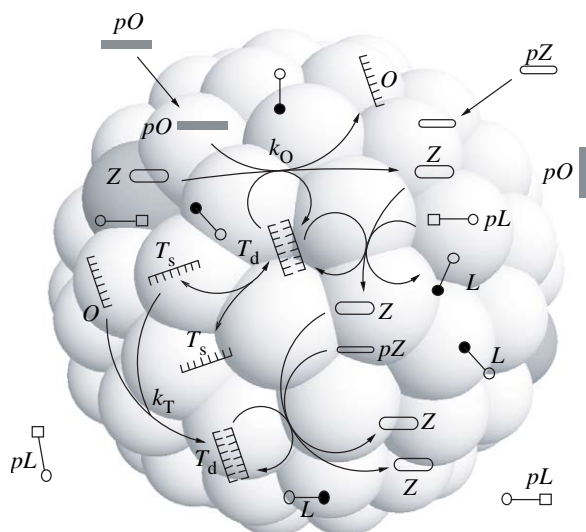
Figure 5. A full protocell model with an explicit metabolism, container and informational molecules. All transformations are coupled through template replication. The molecule Z indicates a sensitizer which is able to take energy from an external (not indicated) source of energy (e.g. light). For key see figure 1.

The system from equations (6.2a)–(6.2d) can be simplified through the change of variable: $x \equiv [T_g]/[A]$, $y \equiv [O_g]/[A]$, $z \equiv [Z_g]/[A]$, variables that are well defined as $[A] \neq 0$ following the reasoning employed in §5. With this change of variables, the system becomes

$$\frac{dx}{dt} = \gamma_T y \sqrt{x} - \gamma_A x^2 z, \tag{6.3a}$$

$$\frac{dy}{dt} = \gamma_0 xz - \gamma_T y \sqrt{x} - \gamma_A xyz, \tag{6.3b}$$

$$\frac{dz}{dt} = \gamma_Z xz - \gamma_A xz^2. \tag{6.3c}$$

Similar to the study of the three-element system, we are interested again in the non-trivial fixed point of equations (6.3a)–(6.3c), considering that the variables x, y and z can take only positive values. The trivial dynamics (no growth) is characterized by $x_0 = 0$ and/or $z_0 = 0$. The non-trivial fixed point of equations (6.3a)–(6.3c) results from $\dot{x} = \dot{y} = \dot{z} = 0$ with $x \neq 0$ and $z \neq 0$ and it is

$$(x_2, y_2, z_2) = \left(\tilde{x}, \frac{\gamma_Z}{\gamma_T} \tilde{x}\sqrt{\tilde{x}}, \frac{\gamma_Z}{\gamma_A} \right), \tag{6.4}$$

$$\text{with} \quad \tilde{x}\sqrt{\tilde{x}} + \frac{\gamma_T}{\gamma_Z} \tilde{x} - \frac{\gamma_O \gamma_T}{\gamma_A \gamma_Z} = 0, \tag{6.5}$$

with the subscript O denoting again the initial conditions, while $\tilde{x}$ is the only positive root of equation (6.5), a conclusion drawn from applying the Routh–Hurwitz criterion to this third-order equation in $\sqrt{x}$. We are not interested in the exact form of $\tilde{x}$, but rather in pointing out its existence and uniqueness in the positive values domain, both features resulting from the Routh–Hurwitz criterion. Through equation (6.2c), this existence leads to $[\dot{Z}_g] \xrightarrow{t \to \infty} \gamma_Z \tilde{x}[Z_g]$, implying an exponential growth of the sensitizer concentration after a sufficiently long time. This result allows us to conclude, with the use of equation

(6.4), that the temporal evolution of all metabolites' concentration is exponential, of the type $\propto \exp(at)$, with $a = \gamma_Z \tilde{x}$.

7. DISCUSSION

There is ongoing debate in the scientific community on the issue of stoichiometric versus catalytic coupling as a necessary dynamics to ensure balanced parallel growth and correct replication of the constituent systems of such a protocell: metabolism, genome and membrane. In the past 20 years, significant experimental efforts have been dedicated to the optimization of self-replicating molecules in order to achieve efficient copying (Paul & Joyce 2004). More precisely, several chemical subsystems must be optimized in order to provide an overall self-replicating system that exhibits efficient auto-catalytic behaviour. The efficiency is directly reflected in the degree of catalysis or the order of reaction, where a value of 1 (exponential growth, recall table 1) is believed to provide the appropriate basis of Darwinian selection: survival of the fittest (Szathmáry & Maynard-Smith 1997).

From the modelling point of view, Gánti's chemoton model (Gánti 1975) is a protocell model incorporating all three necessary subsystems in a very tightly coupled and synchronized interdependence (Mavelli & Ruiz--Mirazo 2007). This model is the basic example of cyclic stoichimetric coupling as the main coordinating factor. In the present work, we extend a catalytically coupled protocell model (Rocheleau *et al.* 2007) and demonstrate that the catalytic coupling of the subsystems is sufficient to lead the whole system into balanced exponential growth. In other words, the metabolism–template–container feedback by itself ensures an overall exponential growth of the system in spite of subexponential or supraexponential local growth of its component subsystems. We believe that this result may have implications for prebiotic and origin of life scenarios, as it is a robust means of providing exponential growth of protocells and thus early Darwinian selection.

In tight connection with the present work, we mention the work of Serra *et al.* (2007). It is also inspired by the Los Alamos Bug and the results of the detailed analytic study therein totally support the coordinated growth of a general class of catalytically coupled protocellular systems.

ENDNOTES

[1] Populations characterized by the same growing exponent.
[2] For the cases $(1-p) < 0$, there exists a time value t_∞ at which the bracket term in the general solution from table 1 becomes zero and due to the negative $(1-p)$ exponent, $x(t_\infty)$ is infinite. Thus, the range of applicability is $0 \leq t \leq t_\infty$.
[3] The value of C_T clearly depends on the value of p. As concentrations have a dimension of $1/V$, V = volume, from equation (4.2a), C_T must have dimensions of $1/(V^{p-1}t)$. As K_t has dimensions of $1/V$ and k_T has dimensions of $(1/Vt)$, we note that the definition of C_T following

equation (4.2*a*) is correct only for $p = 1/2$ case, but needs to be modified for $p \neq 12$.

[4]If all the vesicular protocell's chemistry takes place within the container bilayer itself, then one would still have $V_I \propto S \propto N_L$.

REFERENCES

Bachman, P. A., Luisi, P. L. & Lang, J. 1992 Autocatalytic self-replicating micelles as models for prebiotic structures. *Nature* **357**, 57–59. (doi:10.1038/357057a0)

Bag, B. G. & von Kiedrowski, G. 1996 Templates, autocatalysis and molecular replication. *Pure Appl. Chem.* **68**, 2145–2152.

Eigen, M. 1971 Self-oragization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–526. (doi:10.1007/BF00623322)

Eigen, M. & Schuster, P. 1979 *The hypercycle: a principle of natural self-organization*. Berlin, Germany: Springer.

Gánti, T. 1975 Organization of chemical reactions into dividing and metabolizing units: the chemotons. *Biosystems* **7**, 189–195. (doi:10.1016/0303-2647(75)90057-X)

Gánti, T. 2003 *The principles of life*. Oxford, UK: Oxford University Press.

Kaneko, K. & Yomo, T. 2000 Symbiotic speciation from a single genotype. *Proc. R. Soc. B* **267**, 2367–2373. (doi:10.1098/rspb.2000.1293)

Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. 1996 A self-replicating peptide. *Nature* **382**, 525–528. (doi:10.1038/382525a0)

Luisi, P. L., Ferri, F. & Stano, P. 2006 Approaches to semi-synthetic minimal cells: a review. *Naturwissenschaften* **93**, 1–13. (doi:10.1007/s00114-005-0056-z)

Mavelli, F. 2004 *Theoretical investigations on autopoietic replication mechanisms*. PhD thesis, Eidgenössische Technische Hochschule ETH Zürich.

Mavelli, F. & Ruiz-Mirazo, K. 2007 Stochastic simulations of minimal self-reproducing cellular systems. *Phil. Trans. R. Soc. B* **362**, 1789–1802. (doi:10.1098/rstb.2007.2071)

Monnard, P., Apel, C. L., Kanavarioti, A. & Deamer, D. 2002 Influence of ionic solutes on self-assembly and polymerization processes related to early forms of life: implications for a prebiotic aqueous medium. *Astrobiology* **2**, 213–219. (doi:10.1089/15311070260192237)

Oparin, A. I. 1957 *The origin of life on earth*. London, UK: Oliver and Boyd.

Orgel, L. E. 1992 Molecular replication. *Nature* **358**, 203–209. (doi:10.1038/358203a0)

Paul, N. & Joyce, G. F. 2004 Minimal self-replicating systems. *Curr. Opin. Chem. Biol.* **8**, 634–639. (doi:10.1016/j.cbpa.2004.09.005)

Rasmussen, S., Chen, L., Nilsson, M. & Abe, S. 2003 Bridging nonliving and living matter. *Artif. Life* **9**, 269–316. (doi:10.1162/106454603322392479)

Rasmussen, S., Chen, L., Deamer, D., Krakauer, D. C., Packard, N. H., Stadler, P. F. & Bedau, M. A. 2004 Transition from nonliving to living matter. *Science* **303**, 963–965. (doi:10.1126/science.1093669)

Robertson, A., Sinclair, A. J. & Philp, D. 2000 Minimal self-replicating systems. *Chem. Soc. Rev.* **29**, 141–152. (doi:10.1039/a803602k)

Rocheleau, T., Rasmussen, S., Nielsen, P. E., Jacobi, M. N. & Ziock, H. 2006 Emergence of protocellular growth laws. *Phil. Trans. R. Soc. B* **362**, 1841–1845. (doi:10.1098/rstb.2007.2076)

Segré, D., Ben-Eli, D., Deamer, D. & lancet, D. 2001 The lipid world. *Orig. Life Evol. Biosph.* **31**, 119–145.

Serra, R., Carletti, T. & Poli, I. 2007 Synchronization phenomena in surface-reaction models of protocells. *Artif. Life* **13**, 123–138. (doi:10.1162/artl.2007.13.2.123)

Stadler, B. & Stadler, P. 2003 Molecular replicator dynamics. *Adv. Comp. Syst.* **6**, 47–77. (doi:10.1142/S0219525903000724)

Szathmáry, E. 1991 Simple growth laws and selection consequences. *Trends Ecol. Evol.* **6**, 366–370. (doi:10.1016/0169-5347(91)90228-P)

Szathmáry, E. & Maynard-Smith, J. 1997 From replicators to reproducers: the first major transitions leading to life. *J. Theor. Biol.* **187**, 555–571. (doi:10.1006/jtbi.1996.0389)

Szostak, J. W., Bartel, D. P. & Luisi, P. L. 2001 Synthesizing life. *Nature* **409**, 387–390. (doi:10.1038/35053176)

Takakura, K., Toyota, T. & Sugawara, T. 2003 A novel system of self-reproducing giant vesicles. *J. Am. Chem. Soc.* **125**, 8134–8140. (doi:10.1021/ja029379a)

von Kiedrowski, G. 1986 A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Edn. Engl.* **25**, 932–935.

von Kiedrowski, G., Wlotzka, D.-C. B., Helbing, J., Matzen, D.-C. M. & Jordan, D.-C. S. 1991 Parabolic growth of a hexadeoxynucleotide with a 3′–5′-phosphoramidate linkage. *Angew. Chem. Int. Edn. Engl.* **30**, 423–426. (doi:10.1002/anie.199104231)