Emergence of protocellular growth laws

Tristan Rocheleau\textsuperscript{1,2}, Steen Rasmussen\textsuperscript{1,3,4,*}, Peter E. Nielsen\textsuperscript{4}, Martin N. Jacobi\textsuperscript{5} and Hans Ziock\textsuperscript{1}

1 Self-Organizing Systems, EES-6, Los Alamos National Laboratory, MS-D462, Los Alamos, NM 87545, USA
2 University of California Santa Barbara, Santa Barbara, CA 93106, USA
3 Santa Fe Institute, Santa Fe, NM 87501, USA
4 University of Copenhagen, IMBG, Blegdamsvej 3, 2200 Copenhagen, Denmark
5 Chalmers University of Technology, 41296 Gothenburg, Sweden

Template-directed replication is known to obey a parabolic growth law due to product inhibition (Sievers & Von Kiedrowski 1994 Nature 369, 221; Lee et al. 1996 Nature 382, 525; Varga & Szathmáry 1997 Bull. Math. Biol. 59, 1145). We investigate a template-directed replication with a coupled template catalysed lipid aggregate production as a model of a minimal protocell and show analytically that the autocatalytic template–container feedback ensures balanced exponential replication kinetics; both the genes and the container grow exponentially with the same exponent. The parabolic gene replication does not limit the protocellular growth, and a detailed stoichiometric control of the individual protocell components is not necessary to ensure a balanced gene–container growth as conjectured by various authors (Gánti 2004 Chemoton theory). Our analysis also suggests that the exponential growth of most modern biological systems emerges from the inherent spatial quality of the container replication process as we show analytically how the internal gene and metabolic kinetics determine the cell population’s generation time and not the growth law (Burdet & Kirkwood 1983 J. Theor. Biol. 103, 11–20; Novak et al. 1998 Biophys. Chem. 72, 185–200; Tyson et al. 2003 Curr. Opin. Cell Biol. 15, 221–231). Previous extensive replication reaction kinetic studies have mainly focused on template replication and have not included a coupling to metabolic container dynamics (Stadler et al. 2000 Bull. Math. Biol. 62, 1061–1086; Stadler & Stadler 2003 Adv. Comp. Syst. 6, 47). The reported results extend these investigations. Finally, the coordinated exponential gene–container growth law stemming from catalysis is an encouraging circumstance for the many experimental groups currently engaged in assembling self-replicating minimal artificial cells (Szostak 2001 et al. Nature 409, 387–390; Pohorille & Deamer 2002 Trends Biotech. 20 123–128; Rasmussen et al. 2004 Science 303, 963–965; Szathmáry 2005 Nature 433, 469–470; Luisi et al. 2006 Naturwissenschaften 93, 1–13).\textsuperscript{1}

Keywords: protocell integration; replication; metabolism; kinetics

1. LIPID-ASSOCIATED TEMPLATE-DIRECTED REPLICATION

Our simple protocell model consists of a lipid container (e.g. a large micelle) with amphiphilic replicating protogene molecules and an associated hydrophobic metabolic system as discussed in Rasmussen et al. (2003). However, the details of the coupling between the genes and the container are not important as long as the container defines a localized and restrictive volume or area for the genes (Szostak et al. 2001; Pohorille & Deamer 2002; Rasmussen et al. 2004; Szathmáry 2005; Luisi et al. 2006). In our current minimal representation, only the replicator and container kinetics are modelled explicitly. We can assume that all lipophilic replicators are bound to the exterior surfaces of the lipid aggregates and that both single- and double-stranded templates coexist at these lipid interfaces (figure 1a). For the following arguments, it is only necessary to localize the genes. In our model, we can, for example, assume that the replicator molecules are peptide nucleic acid (Pschl et al. 1998; Nielsen 1999; Mattes & Seitz 2001) oligomers with decorated (lipophilic) backbones. We can further assume that the single- and double-stranded templates are at equilibrium with each other as long as the template replication rate is much slower than the hybridization and dehybridization processes.

The local replication reaction can then be written as

\[ T_s + \alpha O \rightarrow T_d = T_s + T_r, \]  

(1.1)

where \( O \) is a resource replicator substrate molecule and \( \alpha \) is the number of such molecules used in each replication. Denoting these local template concentrations by \( T_s^l \) and \( T_d^l \) for the single- and double-stranded templates, respectively, the local equilibrium condition for the single- and double-stranded templates is then given by

\[ K_r = \frac{(T_s^l)^2}{T_d^l}, \]  

(1.2)

where \( K_r \) defines the equilibrium constant. In the following, we assume that \( K_r \ll 1 \). If \( T_{\text{tot}}^l \) defines the total local template concentration, single- or double
A protocell P is composed of concentration \( T \) (aggregate volume) large micelle) with amphiphilic gene molecules \( G \) at local stranded, we have

\[
T_\text{tot}^i = T_i^1 + 2T_i^2. \tag{1.3}
\]

\( T_i^1 \) and \( T_i^2 \) can now be expressed as a function of \( T_\text{tot}^1 \) by combining equations (1.2) and (1.3),

\[
T_1^1 = -\frac{K_i + \sqrt{K_i^2 + 8K_i^\text{tot}T_\text{tot}^1}}{2} + O_i(K_i), \tag{1.4}
\]

\[
T_2^i = \frac{T_\text{tot}^i}{2} - \frac{T_1^i}{2}
= \frac{T_\text{tot}^i}{2} - \frac{-K_i + \sqrt{K_i^2 + 8K_i^\text{tot}T_\text{tot}^1}}{2}
= \frac{T_\text{tot}^i}{2} \left( \frac{K_i^\text{tot}T_\text{tot}^1}{8} - \frac{O_i(K_i)}{2} \right), \tag{1.5}
\]

where we have used that \( K_i \ll T_\text{tot} \) (\( K_i \ll 1 \)) and \( O_i(K_i) \), \( i = 1, 2 \) is small in terms of \( K_i \).

The left side of equation (1.1) defines the template-directed replication of \( T_i^1 \) using some background precursor template molecules with the local aggregate concentration \( O^i \). The template replication rate equation then becomes

\[
\frac{dT_\text{tot}^1}{dt} = k_{Tf}(O^i)T_1^1 = \frac{1}{\alpha} k_{Tf} O^i T_1^1, \tag{1.6}
\]

where \( f(\cdot) \) is a reaction mechanism dependent, which is \( O^i/\alpha \) in the case where \( \alpha \) resource oligomers are added sequentially to the template one at the time, while \( k_{Tf} \) is the single oligomer addition rate constant. Substituting equation (1.4) into equation (1.6)

\[
\frac{dT_\text{tot}^1}{dt} = k_{Tf}(O^i) \left( \sqrt{K_i^\text{tot}T_\text{tot}^1/2} + O_i(K_i) \right).
\]

where \( O_i(K_i) \) is small, which to the leading order results in parabolic growth

\[
T_\text{tot}^1(t) = \frac{f(O^i)K_i^\text{tot}k_{Tf}^2}{8} + \sqrt{T_\text{tot}^1(0)K_i^\text{tot}f(O^i)k_{Tf} + T_\text{tot}^1(0)}\frac{1}{2}
= \left( \frac{f(O^i)K_i^\text{tot}k_{Tf}}{2\sqrt{2}} + T_\text{tot}^1(0) \right)^2, \tag{1.8}
\]

where \( T_\text{tot}^1(0) \) is the concentration at \( t = 0 \). This result is an example of the well-known growth law for product inhibition in template-directed replication (Sievers & Von Kiedrowski 1994) in an unlimited system.

### 2. TEMPLATE CATALYSED CONTAINER REPRODUCTION

In our model, the aggregate-associated lipid production \( L \) is template catalysed and can be described by

\[
pL + T_d \rightarrow T_d + L, \tag{2.1}
\]

where \( pL \) is the precursor lipid. We assume that the catalysis is dominated by the double-stranded template (Rasmussen et al. 2003). The lipid production rate equation is then

\[
\frac{dL}{dt} = k_L pL^1T_d^1, \tag{2.2}
\]

where \( k_L \) is the lipid production rate constant and \( pL^1 \) is the local (background) lipid precursor concentration.

This simple gene–container coupling can be generalized to explicitly account for a variety of details involved in the metabolic processes as discussed in Munteanu et al. (2007). In this paper, we focus on a detailed discussion of the above minimal gene–container connection.

We assume the average number of lipid molecules per average aggregate volume \( V_A \) to be \( m_0 \). An average aggregate number can, for example, be maintained through a natural micellar size distribution (Rosen 1988). The relative local (dimensionless) aggregate growth rate can now be expressed as

\[
\frac{dV_A}{m_0} = k_L pL^1T_d^1 \frac{V_A}{m_0}. \tag{2.3}
\]

Let \( A \) be the global aggregate concentration. Since each aggregate is assumed to divide following growth in order to maintain an average of \( m_0 \) lipid molecules per aggregate, the growth in the global aggregate concentration due to the local lipid production in the aggregates can be expressed as

\[
\frac{dA}{dt} = AV_A \frac{dL}{m_0} \frac{dV_A}{m_0} = k_L pL^1T_d^1 \frac{AV_A}{m_0}. \tag{2.4}
\]

In the above expression, we have implicitly assumed that the volume of the templates and its change are small (constant) compared with the volume of lipid in
an aggregate. We can now transform the local aggregate concentration of templates into the corresponding global concentrations using the simple relation between local and global template concentrations

\[ T^g = T^l A V_A, \] (2.5)

noting that \( T^l \) is measured in template molecules per lipid aggregate volume \( V_A \) and \( A \) is measured in the number of lipid aggregates (of average size \( V_A \)) per system volume. Substituting equation (2.5) into equation (2.4) and taking into account that \( T_{tot} = 2T_{db} \) we obtain

\[ \frac{dA}{dt} = k_L p^l \frac{T^g}{m_0} T^g \gamma_A T^g_{tot}, \] (2.6)

where

\[ \gamma_A = \frac{k_L p^l}{2m_0}. \] (2.7)

It should be noted that we can just as well express the local concentrations in molecules per aggregate, in molecules per aggregate volume using an average aggregate volume \( V_A \) or in molecules per aggregate surface area \( S_A \).

3. GLOBAL TEMPLATE REPRODUCTION

The growth equation for the local total template concentration \( T_{tot} \) corresponding to equation (1.6) is given by

\[ \frac{dT^g_{tot}}{dt} = k_T f(O^l) T^l \frac{dA}{dt} T^g_{tot} - \frac{dA}{dt} T^g_{tot} A, \] (3.1)

where the second term stems from the effect the volume change has on the local concentration. Since equation (3.1) is central to the argument, we discuss it in detail. The question that needs to be answered is what happens to the concentrations when one goes from a well-stirred solution to many small containers. This has to do with the origin of the second term in equation (3.1). The volume of the container grows with time and hence the concentration can change, and decreases with increasing container volume as \( A \) is measured in aggregates of average size \( V_A \) per litre. The volume growth of the container must be equal to the growth in the (global) concentration of the aggregates. One also needs a normalization factor to go from the change in container volume to the change in template concentration. All together, this yields a minus sign (the concentration goes down with increasing volume) multiplied by the ratio of the concentration of templates divided by the concentration of aggregates.

Now, since equation (1.2) holds for the local concentration (not the global), we can use the approximate form of equation (1.4) to rewrite the growth equation as

\[ \frac{dT^g_{tot}}{dt} = k_T f(O^l) \frac{\sqrt{K_T T^l}}{2} - \frac{dA}{dt} \frac{T^l}{A} \frac{T^g_{tot}}{A}. \] (3.2)

To rewrite the growth in terms of global concentrations, we multiply the growth equation (3.2) by \( A \), and get

\[ \frac{dT^g_{tot}}{dt} = k_T f(O^l) \frac{\sqrt{K_T T^l}}{2} A - \frac{dA}{dt} \frac{T^l}{A} T^g_{tot}. \] (3.3)

Note that the volume compensation term vanishes as it should. We then use the relation \( T^g = T^l A V_A \) to rewrite the growth in terms of global concentrations

\[ \frac{dT^g_{tot}}{dt} = \gamma_T \sqrt{T^g_{tot} A}. \] (3.4)

where

\[ \gamma_T = k_T f(O^l) \sqrt{K_T V_A \frac{2}{2}}. \] (3.5)

4. COUPLED TEMPLATE AND CONTAINER REPRODUCTION

The coupled growth equations (2.6) and (3.4) can be solved by separation of variables, which can be used to rewrite equation (3.4) where we are seeking the solution for large \( T^g_{tot} \) values. We obtain (Appendix A) the solution,

\[ T^g_{tot}(t) = \frac{(\exp \left( \frac{1}{2} a t + D_0 \right) - b)^{2/3}}{a^{2/3}} \sim \exp(a t), \] (4.1)

where

\[ a = \gamma_T \gamma_A \] (4.2)

and

\[ b = \frac{\gamma_T^2 C_0}{3 \gamma_T^2}, \quad C_0 = A(0)^{2/3} - \frac{\gamma_A}{\gamma_T} T^g_{tot}(0)^{3/2}. \] (4.3)

Substituting equation (4.1) into equation (2.6) yields

\[ \frac{dA}{dt} \sim \gamma_A \exp(a t), \] (4.4)

which means that

\[ A(t) \sim \exp(a t) \sim T^g_{tot}(t). \] (4.5)

Thus, to a leading order, both the coupled template and the aggregate grow exponentially with the same exponent \( a \). As equations (2.7) and (3.5) are substituted into equation (4.2), we note that the exponent in a direct manner depends on the template and lipid production rate constants, \( k_T \) and \( k_L \), the local background precursor concentration, \( p_L \), the background template substrate concentration and the order of the template reaction, \( f(O^l) \), the template duplex equilibrium constant, \( K_t \), and the average number of lipids per aggregate, \( m_0 \).

\[ a = \left( \frac{1}{4} k_T^2 k_L p_L f(O^l)^2 K_T V_A \right)^{1/3}. \] (4.6)

In figure 2, a simulation of the original kinetic equations (2.6) and (3.4) illustrates these findings. It should be noted that at least one gene duplex must be present in each aggregate for the derived kinetic growth equations to be valid, which is ensured if \( \gamma_A < \gamma_T \).

5. DISCUSSION

A kinetic analysis of a minimal protocell model shows that the local parabolic template growth law is compensated by parallel aggregate division dynamics, which results in an overall exponential growth for the

\[ \text{Phil. Trans. R. Soc. B (2007)} \]
Thus, a detailed stoichiometric growth control of the individual protocell components is not necessary as conjectured, for example, by Gánti (2004). A coupled autocatalytic feedback between the template and the aggregate replication processes generates a coordinated growth of the main components. This result is an extension of the exponential growth law in the low template concentration limit as discussed by Stadler (Stadler & Stadler 2003) and others. As long as the effective average local replicator concentration is kept constant, in this case by growing the system’s total lipid volume, exponential growth is also possible. A discussion of the detailed growth dynamics in modern cells can be found in e.g. Burdett & Kirkwood (1983); Novak et al. (1998) and Tyson et al. (2003).

We thank the Protocell Assembly team (http://protocells.lanl.gov/team), Camille S.-O. Attolini, Jerzy Maselko and Andreea Munteanu for their critical feedback on the calculations and earlier versions of this manuscript. We also thank Eörs Szathmáry for pointing us to early work on intracellular growth kinetics. This work is supported in part by the Los Alamos National Laboratory LDRD-DR grant on ‘Protocell Assembly’ (PAs) and by the European Commission’s Sixth Framework project on ‘Programmable Artificial Cell Evolution’ (PACE).

ENDNOTES

1It should be noted that the main result presented in this paper can be generalized as demonstrated in Munteanu et al. (2007).

2It should be noted that an exponential amplification is also possible in a naked template-directed replication process if the resulting new template is modified in the ligation process to lower its hybridization time.

APPENDIX A

The coupled growth equations (2.6) and (3.4) can be solved by separation of variables, i.e. by dividing equation (3.4) by equation (2.6)

\[
\frac{dT^g_{\text{tot}}}{dA} = \frac{\gamma_T}{\gamma_A} \sqrt{A}\cdot \frac{\sqrt{T^g_{\text{tot}}}}{T^g_{\text{tot}}}.
\]  
(A 1)

Note that equation (A 1) is a special case of a Bernoulli differential equation. From equation (A 1), we obtain

\[
(A(t))^{3/2} = \frac{\gamma_A}{\gamma_T} (T^g_{\text{tot}}(t))^{3/2} + C_0,
\]  
(A 2)

where \(C_0 = A(0)^{3/2} - \frac{\gamma_A}{\gamma_T} (T^g_{\text{tot}}(0))^{3/2}\). We use this to rewrite equation (3.4) as

\[
\frac{dT^g_{\text{tot}}}{dt} = \gamma_T \sqrt{T^g_{\text{tot}}} \left(\frac{\gamma_A}{\gamma_T} (T^g_{\text{tot}})^{3/2} + C_0\right)^{1/3}
\]

\[
= \gamma_T \sqrt{T^g_{\text{tot}}} \left(\frac{\gamma_A}{\gamma_T} \left(\left(\frac{\gamma_A}{\gamma_T} (T^g_{\text{tot}})^{3/2} + C_0\right)^{1/3} \right)^{3/2} + \frac{C_0 \gamma_T^{3/2}}{3 \gamma_A} T^g_{\text{tot}} + O\left((T^g_{\text{tot}})^{-3/2}\right)\right),
\]

(A 3)

where we have expanded the expression for large \(T^g_{\text{tot}}\) relative to \(C_0\). Ignoring the last term, which is small in order \(T^g_{\text{tot}}^{-3/2}\), the simplified expression has the solution

\[
T^g_{\text{tot}}(t) = \left(\exp\left(\frac{1}{2} at + D_0\right) - b\right)^{2/3} \sim \exp(at),
\]  
(A 4)
Emergence of protocellular growth laws  T. Rocheleau et al.  1845

to the leading order in $T_{\text{tot}}$, where $a = \gamma_2^{1/3} \gamma_3^{2/3}$, $b = \left(\gamma_2^{2/3} C_0/3\right) \gamma_1^{2/3}$ and $D_0 = \ln(b + a T_{\text{tot}}(0)^{3/2})$. Substituting equation (4.1) into equation (2.6) yields

$$\frac{dA}{dt} \sim \gamma_4 \exp(at),$$

(A 5)

which means that

$$A(t) \sim \exp(at) \sim T_{\text{tot}}^6(t).$$

(A 6)

Thus, to leading order, the coupled template and aggregate grow exponentially with the same exponent. Simulation of equations (2.6) and (3.4) verifies these findings (figure 2).

REFERENCES


