Synthetic Protocell Biology: From Reproduction to Computation

Ricard V. Solé
Andreea Munteanu
Carlos Rodriguez-Caso
Javier Macia

SFI WORKING PAPER: 2006-09-033

SFI Working Papers contain accounts of scientific work of the authors and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
Synthetic Protocell Biology: from reproduction to computation

Ricard V. Solé,1,2 Andreea Munteanu,1 Carlos Rodriguez-Caso,1 and Javier Macía1
1ICREA-Complex Systems Lab, Universitat Pompeu Fabra (GRIB), Dr Aiguader 88, 08003 Barcelona, Spain
2Santa Fe Institute, 1399 Hyde Park Road, Santa Fe NM 87501, USA

Cells are the building blocks of biological complexity. They are complex systems sustained by the coordinated cooperative dynamics of several biochemical networks. Their replication, adaptation and computational features emerge as a consequence of appropriate molecular feedbacks that somehow define what life is. As the last decades have brought the transition from the description-driven biology to the synthesis-driven biology, one great challenge shared by both the fields of bioengineering and origin-of-life is to find the appropriate conditions under which living cellular structures can effectively emerge and persist. Here we review current knowledge (both theoretical and experimental) on possible scenarios of artificial cell design and their future challenges.

Keywords: cells, cell cycle, cell membrane, metabolism, information, synthetic biology

I. INTRODUCTION

Cellular life cannot be described in terms of only DNA (or any other information-carrying molecule) nor as metabolism or as compartment (cell membrane) alone. Cellular life emerges from the coupling among these three components. A container is a prerequisite of biological organisation in order to at least confine reactions in a limited space, where interactions are more likely to occur (Deamer et al., 2002). It also provides a spatial location able to effectively facilitate division of labour among reactions. Moreover, in modern cells, the membrane is an active cell component, channelling nutrients in and waste products out of the cell by means of specialised transport catalysts (Pohorille et al., 2005). A metabolism (Smith & Morowitz, 2004) provides the source of non-equilibrium and a mean of energy storage required in order to build and maintain cellular components. It is also required to allow cell growth to occur and eventually force splitting into two different (but similar) copies.

If cells must adapt to a changing environment, information carriers will be also needed, as well as their coupling with metabolic reactions. The fundamental problem, of course, is how to obtain the coupling in such a way that self-reproduction of identical structures is achieved and self-maintained. Even though the scientific community has reach a consensus on the requirements and properties of a minimal living system (Deamer, 2005; Pohorille & New, 2001), the materialisation of this vision into a concrete laboratory prototype is still incomplete.

In this review we explore some of the methods, theories and perspectives related to the goal of constructing simple artificial cellular systems. The research program to which this special number is dedicated and which we name synthetic protocell biology (SPB) aims at the construction of a chemical life-like ensemble in the form of an artificial cell system able to self-maintain, self-reproduce and potentially evolve. The article is organised as follows: as SPB can be considered as a field intersecting synthetic biology, which is a general framework encompassing systems biology and bioengineering, we shall place the for-
distinctive feature of synthetic biology seems to regard the emphasis on design and testing via simulation of new living biochemical systems endowed with complex behaviour, followed by their experimental implementation. From this point of view, both synthetic biology and SPB follow the same course of action.

Two main research branches can be delimited within the field of synthetic biology. A first one concerns the SPB: the assembly of synthetic units into chemical systems endowed with biological properties, that is reproduction, inheritance and evolution (Pohorille & Deamer, 2002). The second branch aims at assembling biological units extracted from living systems and obtaining a modified or even an improved version of a (an existing) biological system. This second branch confers a clear engineering feature to synthetic biology by dealing with the creation and rewriting of genetic circuits using building blocks (Endy, 2005).

One of the main goals pursued by both research branches consists in constructing, modifying and using biological mechanisms into performing desired functions, in other words, in obtaining a programmable plug-in genetic device or biological entity (Kobayashi et al., 2004). Either by a rational complete design (Weiss et al., 2003) or by directed evolution techniques (using combinatorial synthesis: Yokobayashi et al. 2002), programming artificially designed living protocells appears as the world-wide synthetic biology objective. In addition to programmable features, SPB also aims at integrating a property unique to living matter: evolvability. Even though there still is a long way to go before achieving this objective, there is active ongoing research on the evolutionary potential of protocellular and prebiotic structures (Szathmary et al., 2005; Yokobayashi et al., 2003). Moreover, the concept of computation is intimately connected to both evolvability and programmability features and latest studies point to it as a fundamental characteristic of biological systems (Fernández & Solé, 2003). We shall discuss momentarily these concepts and the relationship between them in the framework of SPB.

III. BUILDING PROTOCELLS

The major works dedicated to SPB are organised in Table I. Even though SPB is a relatively young research field, it is related to as well as a continuation of several research fields and thus its frontiers are not clearly established. For this reason, the task of choosing among the most significant and relevant results obtained so far is extremely difficult. As a general rule, we chose to cite mainly review papers that incorporate direct citations of particular and detailed works.

Two fundamental approaches have been considered towards the synthesis of a protocell (fig. 1). The first class, so called top-down approach (Luisi et al., 2002) involves the creation of a minimal cell by means of reducing the genome of modern cells. Since numerous genes are involved in cell-cell communication while others have been shown to be non-essential to cell functioning, it was early suggested that it would be possible to reduce genome complexity to a minimal set of genes able to sustain (under given external conditions) cell life and reproduction. In this context, using the parasitic bacterium Mycoplasma genitalium as a case study, Mushgeian & Koonin (1996) showed that approximately 256 genes seem to be needed to build a minimal gene set that is necessary and sufficient to sustain the existence of a modern-type cell. No matter how small, cell genomes must contain all the information necessary for the cells to perform essential (housekeeping) functions allowing them to maintain metabolic homeostasis, reproduce and evolve, the three main properties of living cells. Moya and co-workers (Gil et al., 2004) have also studied this problem using both genome comparison and computational modelling approaches, further reducing the list of essential genes to only 206 (see also Gabaldon et al., 2006).

In this paper we restrict our attention to the bottom-up approach. Contrary to the top-down approach, it starts from scratch: a life-like entity is build by self-assembling of molecular components (Rasmussen et al., 2003). These can be of biological nature or instead completely ad hoc chemical components. In both cases, a compartment is required, formed of some type of amphiphilic molecules characterised by a natural tendency to make aggregates to which further cell components can join and couple. Plymers that mimic lipid amphiphility can also assemble into vesicles, offering a much wider spectrum of properties (Discher & Eisenberg, 2002).

To detail even further, Luisi et al. (2006) suggest also the term reconstruction as the intermediate between the bottom-up and the top-down approaches. This term refers to the encapsulation of nucleic acids and enzymes in liposomes defining thus a “semi-artificial cell” (Po-
horille & Deamer, 2002) whose self-replication is again sought in the experimental (Oberholzer et al., 1995) and theoretical works (Szostak et al., 2001).

As indicated in fig. 1, the final outcome of both approaches does not need to be a self-replicating, evolving cell. The artificial cell (Acell) can replicate but not evolve, or it might even be unable to replicate itself: such simple self-maintained Acell would be able to persist under given conditions, exchanging energy and matter from its surroundings without growing. Although this seems a fairly limited situation, it is actually an interesting one: one could design or lead into self-assembling a protocell able to perform some given functions or computations in predefined ways (Pohorille & Deamer, 2002). These two major concepts, self-assembling and self-replication, pervade the field of SPB and origin-of-life and we shall discuss them below in more detail.

A. A self-replicating machine

Since building a protocell using a bottom-up approach is not limited to biological, evolved ingredients, we should ask ourselves if very different combinations of the previous three ingredients (or none of them) might lead to life-like entities. To answer this question requires formulating the problem under a theoretical perspective. Such view was taken by the Hungarian mathematician John von Neumann in the 1940’s (von Neumann, 1966). In his seminal work, von Neumann considered the logical conditions under which an abstract - but embodied - automaton would be able to reproduce itself. The approach was fully computational: the reproducing system was viewed as a machine equipped with building blocks and instructions. The solution found was innovative and visionary. A basic scheme is shown in fig. 2. Here the basic building blocks are indicated, including:

1. The constructor (A), able to build a physical, new system by using the available raw material in the surroundings.

2. The blueprint or instructions containing information on what has to be performed by the constructor.

3. A duplicator (B) which takes the instructions and duplicates them accurately.

4. The controller (C) required to guarantee that the whole process takes place in some well-defined sequence.

Although not explicitly said in the previous description, the automaton envisioned by von Neumann had a physical embodiment: he specifically thought on how to build a physical system able to perform the whole replication cycle. In other words, von Neumann’s system was a machine. He designed a system which should have 29 states and estimated that on the order of $10^3$ elements would be required. However, smaller versions of the system have been obtained even at the hardware level (Reeves et al., 2000).

One great implication of von Neumann’s solution was the need for a copy mechanism for the instruction set. If no such mechanism were present, an infinite recurrence problem would arise: the instructions would have to contain a replica of themselves which would be passed to the new machine. Since he investigated the required logic of replication, he was not interested, nor did he have the necessary tools, in building a working machine at the bio-chemical or genetic level. Remember that at his time DNA had not yet been discovered as nature’s genetic material. The fact that von Neumann’s approach was so close to current cell organisation is revealing. It suggests that universal principles might pervade the ways information, metabolism and container need to be coupled. We can identify the components of the automaton with those of living cells as follows: (a) the instruction set is the DNA molecule; (b) the duplicator is provided by the DNA polymerase and other components of the cell’s replication machinery; (c) the constructor corresponds to the RNA polymerase and the translation machinery (allowing proteins to be formed) and (d) the controller is nothing but the regulation of transcription and translation.

The road initiated by von Neumann has been followed by many other researchers over the last decades of the 20th century. As molecular biology advanced, new tools and methods allowed to consider a different approach to the problem of self-replication: replacing machines by biological components. In this context, several important advances were obtained within the context of self-replicating molecules, both conceptually (Szathmáry & Maynard-Smith, 1997; von Kiedrowski, 1993) and experimentally (see Paul & Joyce, 2004, and references therein). These include different small-sized molecules able to display a closed replication cycle.
FIG. 3 A schematic self-reproduction cycle of a protocell, requiring growth (G) and replication (R) phases in order to be completed. When dealing with a nano-scale scenario, both internal and external noisy fluctuations ($\xi$) are expected to affect the reliability of the whole process.

The machine’s self-reproduction envisioned by von Neumann has an equivalent picture within the context of protocell replication. The question here is: what are the conditions allowing a simple artificial protocell to reach reliable reproduction? von Neumann’s picture includes two key components of a complex adaptive system able to process information: hardware and software. In modern cells, software is carried by DNA, whereas proteins play the role of cellular hardware. But in order to achieve reproduction, no software is required (Dyson, 1999); if the appropriate mechanisms of molecular assembly are in place and the system is able to spontaneously grow out from equilibrium, replication can occur without the designed concurrence of all the von Neumann’s components. In this context, the physical and chemical properties of molecules are able to define a replicating entity without using the machine-like picture of sequential operations. In this self-organised picture, the basic cell cycle includes two steps: growth and division, summarised in fig. 3. Both steps need the presence of non-equilibrium conditions unless some externally driven mechanism (such as shear forces) is present. And they are affected by both internal and external fluctuations, here indicated as noise sources ($\xi$). Basic physical and chemical constraints are the main players in this process, connected with the stability of membrane shape. To a large extent, it is the active breaking of such stability towards non-spherical vesicle shapes what predates the problem of protocell reproduction. This has been explored by a number of authors (Božič & Svetina, 2004; Du et al., 2006; Jung et al., 2001; Seifert et al., 1991; Svetina & Žekš, 1989) and is formulated in terms of the bending energy function $\mathcal{H}_b$ associated to the closed vesicle. If we indicate as $\mathcal{H}_b[S]$ the free energy density at a given point $S$ on the surface $S$ of the lipid bilayer, the total bending energy will be the integral over the cell’s surface $S$:

$$\mathcal{H}_b = \int_S \mathcal{H}_b[S] dS$$

(1)

In its simplest form and considering low temperatures (i.e. thermal fluctuations are ignored) we have

$$\mathcal{H}_b = \int_S \frac{\kappa[S]}{2} (C(S) - C_0(S))^2 dS$$

(2)

where $\kappa[S]$ is the bending modulus and $C(S) - C_0(S)$ is the mean curvature of the vesicle surface at $S$. The result of the minimisation of such energy function, i.e. the solutions of

$$\delta \mathcal{H}_b = 0$$

(3)

allows to explain a number of basic vesicle shapes, from red blood cells (Du et al., 2006; Jülicher, 1996; Pirotto & Mavelli, 2004; Seifert et al., 1991; Waugh, 1996; Zhongcan & Helfrich, 1987) to potential conditions for self-reproduction (Božič & Svetina, 2004; Svetina & Žekš, 1989). In a more general context, it can be easily shown that energy configurations forbid spontaneous splitting of a spherical vesicle to occur under the absence of other energy sources (Solé et al., 2006). Below we will explore some of the potential scenarios able to allow cell division to take place. Evolution of cell division and other membrane remodelling mechanisms has led to sophisticated mechanisms of organization and self-assembly that make them more and more independent of external fluctuations (Shapiro & Benenson, 2005). However, much simpler protocells face a rather different situation, where the presence of noise can be an inevitable component to be taken into account, particularly when dealing with nanoscale systems (Fellermann and Solé, 2006).
In von Neumann’s picture, self-reproduction was based on a deterministic sequence of events where the software and hardware of the machine interacted in a predictable way. When dealing with wet machines, we must take into account the spontaneous contribution of self-assembly processes and the constraints derived from them. We can also move away from the machine carrying instructions to be copied. Here such information is embedded in the container and its interactions with metabolism and information, but the latter can also be removed from the system. Nevertheless, ongoing experimental efforts have explicitly considered von Neumann’s approach by using a small set of genes to be enclosed within a vesicle and truly acting as a software system able to sequentially control a chain of events eventually leading to cell replication (Noireaux et al., 2005).

B. Self-assembly vs design

Most of these potentially feasible levels of protocell complexity are linked to vesicle containers. Thus, special attention has been dedicated to exploring the behaviour of these self-organised structures. Three decades ago, a first crucial step towards the bottom-up synthesis of life consisted in the experiments on self-assembly phenomena leading to microscopic gel structures: Oparin’s coacervates (Oparin & Gladilin, 1980) and Fox’s proetoid microspheres. Early continuations of these scenarios of self-assembling prebiotic containers have focused on surfactant assemblies into micelles and vesicles. By now it is clear that the lipid vesicles are the meeting point of the top-down and bottom-up approaches, as the idea of cellular life evolving from within a compartment is universally accepted (Luìsi et al., 1999). More precisely, experimental works (Monnard & Deamer, 2002) complemented by computer simulations of self-assembly of amphiphiles (Fellermann & Solé, 2006) made the object of study in the bottom-up approach, while the top-down and reconstruction studies focused on the more biological-like membrane structure called liposomes (Oberholzer & Luìsi, 2002).

Liposomes are lipid vesicles (bilayers) typically prepared from phospholipids, the components of today’s cells’ membrane. Besides being improbable if not impossible that such molecules existed in prebiotic environment, they confer a low permeability to the constituting membrane and thus a high resistance to the uptake of nutrients (Deamer et al., 2002). Even in the absence of the evolved transport mechanisms of today’s living cells, there are solutions to improve liposomes’ permeability (Monnard & Deamer, 2001). From the point of view of both origin of life and life synthesis, the use of fatty acids (surfactants) in membrane structures instead of phospholipids allows a higher permeability to ionic solutes and thus seems more plausible and efficient.

In order to achieve the desired behaviour, researchers have also used a number of basic building blocks in most experimental designs. Some of them are shown in fig. 4. These include transmembrane proteins able to facilitate the active flow of precursors, special enzymes and polymerases taken from viruses and bacteria and other well-characterised molecules. These are of course only a few items from a potentially huge universe of possible molecular structures, and the ongoing progress of synthetic biology, with interacting sets of molecules defining oscillators or switches, will help designing protocells able to perform complex functional tasks.

C. Building scenarios

As previously mentioned, the end point of SPB is the building of artificial cells able to behave in a life-like manner. This includes self-reproduction, self-maintenance and evolvability. But different intermediate stages can be also relevant, even if they do not incorporate all the previous ingredients. Before we present the different approaches taken, a tentative list of artificial protocell types can be defined as follows:

1. **Self-maintaining protocells**: in this type of synthetic protocell, neither growth nor replication would take place (e.g. Zepik et al. 2001). However, active transfer of energy and matter can occur by means of appropriate energy-transducing mechanisms (such as transmembrane proteins pumping protons) – Table 1(5). Although this cell is unable to replicate itself, it might have desirable properties such as acting as a nanomachine able to process
<table>
<thead>
<tr>
<th>Approach</th>
<th>Container</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bottom-Up</strong></td>
<td></td>
</tr>
<tr>
<td>(1) Lipid world</td>
<td>Segré et al. (2001)</td>
</tr>
<tr>
<td>(2) Self-catalytic lipid aggregate formation. Theory &amp; Simulations</td>
<td>Mavelli &amp; Luisi (1996)</td>
</tr>
<tr>
<td>(3) Self-catalytic lipid aggregate formation. Experiments</td>
<td>Bachman et al. (1992); Berclaz et al. (2001a, b); Chen &amp; Szostak (2004); Hanczyc &amp; Szostak (2004); Takakura &amp; Sugawara (2004); Takakura et al. (2003)</td>
</tr>
<tr>
<td>(4) Self-catalytic lipid aggregate formation. Simulations</td>
<td>Chen &amp; Kim (2004); Lyubartsev (2005); Nilsson et al. (2003); Noguchi &amp; Takasu (2002)</td>
</tr>
<tr>
<td>(5) Increasing bilayer permeability</td>
<td>Monnard &amp; Deamer (2001)</td>
</tr>
<tr>
<td>&gt; channels: Noireaux &amp; Libchaber (2004); Pohorille et al. (2005)</td>
<td></td>
</tr>
<tr>
<td>&gt; microtubulation: Roux et al. (2002)</td>
<td></td>
</tr>
<tr>
<td>&gt; alternating electric field: Fischer et al. (2002)</td>
<td></td>
</tr>
<tr>
<td>(6) Thermodynamics of self-assembling of lipid aggregates</td>
<td>Mayer et al. (1997)</td>
</tr>
<tr>
<td>(7) Fusion of liposomes-vesicles</td>
<td>Pantazatos &amp; MacDonald (1999)</td>
</tr>
<tr>
<td>(8) Photoinduced vesicle formation</td>
<td>Veronese et al. (1998)</td>
</tr>
<tr>
<td><strong>Biochemical reactions in container.</strong></td>
<td></td>
</tr>
<tr>
<td>(9) Conceptual framework of macromolecules encapsulation</td>
<td>Szostak et al. (2001)</td>
</tr>
<tr>
<td>(10) Protein expression in liposomes</td>
<td>Oberholzer et al. (1999)</td>
</tr>
<tr>
<td>(11) Increased efficiency of gene expression (transcription of T7 DNA) in liposomes</td>
<td>Nomura et al. (2003); Tsumoto et al. (2001)</td>
</tr>
<tr>
<td>(12) Two-stage genetic network encapsulated in liposomes</td>
<td>Ishikawa et al. (2004)</td>
</tr>
<tr>
<td>(13) Enzyme-driven vesicle replication</td>
<td>Hanczyc et al. (2003); Szostak et al. (2001); Walde et al. (1994)</td>
</tr>
<tr>
<td>(14) Coupled RNA-template and vesicle replication. Experiments</td>
<td>Chen et al. (2004); Oberholzer et al. (1995)</td>
</tr>
<tr>
<td>(15) Homeostatic behaviour</td>
<td>Zepik et al. (2001)</td>
</tr>
<tr>
<td><strong>Bottom-Up</strong></td>
<td></td>
</tr>
<tr>
<td>(17) Coordinated shell-core growth. Simulations</td>
<td>Munteanu et al. (2006); Serra et al. (2006 (In press)</td>
</tr>
<tr>
<td><strong>Information (I) / Metabolism (M) / Container (C)</strong></td>
<td></td>
</tr>
<tr>
<td>(20) CI System</td>
<td>Szostak et al. (2001), Fontanari et al. (2006)</td>
</tr>
<tr>
<td>(21) MCI System. Simulations</td>
<td>Fernando &amp; Di Paolo (2004); Munteanu et al. (2006); Zintzaras et al. (2002)</td>
</tr>
<tr>
<td><strong>Evolutionary potential</strong></td>
<td></td>
</tr>
<tr>
<td>(22) General implications of autocatalytic subsystems</td>
<td>Kaneko (2005)</td>
</tr>
<tr>
<td>(23) Minimal protocell as unit of evolution</td>
<td>Száthmáry et al. (2005)</td>
</tr>
<tr>
<td>(24) Evolving replicators in limited dispersion</td>
<td>Szabó et al. (2002)</td>
</tr>
<tr>
<td><strong>Top-Down Minimal cell</strong></td>
<td></td>
</tr>
<tr>
<td>(25) Generalised kinetic model of E Coli metabolic mapping</td>
<td>Browning &amp; Shuler (2001)</td>
</tr>
<tr>
<td>(26) Conceptual framework of minimal gene set</td>
<td>Gil et al. (2004); Koonin (2000); Luisi et al. (2006, 2002)</td>
</tr>
<tr>
<td>(27) Experiments on minimal genome</td>
<td>Hutchinson et al. (1999); Mushgian &amp; Koonin (1996)</td>
</tr>
<tr>
<td><strong>Computational potential</strong></td>
<td></td>
</tr>
<tr>
<td>(29) Protein molecules as computational elements</td>
<td>Bray (1995)</td>
</tr>
<tr>
<td>(30) Design of genetic circuits</td>
<td>François &amp; Hakim (2004); Savageau (2001)</td>
</tr>
<tr>
<td>(31) Programmable functions</td>
<td>You et al. (2004)</td>
</tr>
</tbody>
</table>
matter or information under given external conditions. A biosensor or a drug delivery system can be potentially implemented at this level.

2. Growing vesicles: due to active transformations of external precursors, growth can take place over some transient time. These systems eventually stop growing (once an equilibrium between internal metabolism and matter exchanges is reached), but can be helpful in exploring through both experiments and simulations the mechanisms that allow self-aggregates and artificially designed vesicles to grow – (Chen & Szostak, 2004) Table I(1-8). This system can also be a first step towards cell division by using extrusion mechanisms through membrane pores – Table I(13).

3. Replicating vesicles: if vesicles or micelles are driven out from equilibrium and keep growing, they can eventually reach unstable configurations favouring membrane asymmetries, deformation and eventually division (Hanczyc & Szostak, 2004).

The simplest scenario does not include information-carrying molecules and thus does not consider evolution. However, in spite of the limitations imposed by evolution-free systems, it might actually be important in a number of applications (such as biomedicine) where evolvability is actually something to be avoided.

4. Replicating protocells: “reconstruction” approach is probably one of the most active directions in SPB, where macromolecules are encapsulated in vesicles or liposomes and catalyse metabolic functions necessary in the life cycle of the protocell. One such example is Szostak’s theoretical minimal RNA cell, not yet implemented in laboratory in the exact composition as prescribed by Szostak et al. (2001). However, different versions of the RNA cell have been created in the laboratory (Bartel & Unrav, 1999; Chen et al., 2004). In the long list of experimental works, an important place is held by the works of Walde et al. (1994a) and Oberholzer et al.
(1995) showing “core-shell” replication in designed protocells. This coordinated replication of both container and metabolites (and/or genetic material) is necessary in order to avoid death by dilution after several generations of protocells. Different from the reconstruction approach, the Los Alamos Bug model (Rasmussen et al., 2003) and versions of it, based on the self-assembly approach, have been proved to fulfill this condition, an emergent property of the catalytic coupling of protocell’s sub-systems – Table I (17). Opposed to these catalysis-based protocell models, a stoichiometric model of a protocell, the chemoton model (Gánti, 2003) accomplishes the coordinated growth of all its components by means of a precise imposed stoichiometry in the transformation nutrients-metabolites-waste.

5. Evolving protocells: if the macromolecular organisation of the SPB includes an information carrier coupled to metabolic and container dynamics, the whole assembly can experience evolutionary changes and Darwinian selection (Szathmáry et al., 2005). For example, in the stoichiometrically-coupled chemoton model (Gánti, 2003), the information is carried by polymers of a given length, and thus changes in the number of constituting monomers can induce changes in the efficiency of protocell dynamics and thus a genotype-phenotype mapping.

IV. PROTOCELL MODELS

In the context of the scenarios previously mentioned, several models of synthetic protocells can be envisioned. In this section we enumerate some possible examples of systems that allow cell self-replication by using different potential mechanisms. Some of these systems have been explored from the mathematical point of view.

1. Enzyme-driven, information-free protocells: based on an early suggestion by the Russian biomathematician Nicolas Rashevsky, it has been shown that the stability of a spherical closed membrane can be lost by providing an appropriate set of metabolic reactions (Rashevsky, 1960). Specifically, Rashevsky conjectured that a vesicle having enzymes allowing metabolic reactions to occur inside the compartment could experience a destabilisation eventually leading to cell division. It has been recently shown that this is the case for a very simple model involving, for simplicity, two enzyme molecules, placed at the two opposite poles of the cell (fig. 5a). If these enzymes catalyse a reaction transforming a given precursor (R in figure 5a) into a new molecule G, then close to the location of each enzyme, the concentration of G would increase and eventually trigger a heterogeneous pressure distribution along the membrane (Macía & Solé, 2006a). Additionally, surrounding lipid molecules L become incorporated into the vesicle as membrane bounded molecules (here indicated as $L_m$). Since the process is necessarily linked to enzymatic activity, the splitting is a single event, not able to be repeated.

2. Turing-like protocells: these model protocells introduce a very simple coupling between a reaction-diffusion system defining a metabolism and membrane dynamics (fig. 5b). In this scenario there is no need for spatially localised enzymes. Instead, externally provided precursors $R_1, R_2$ are supplied, entering the membrane and being transformed into new molecules through a set of simple reactions. The reactions inside and outside the cell are represented by a set of n reaction-diffusion (RD) equations (Murray, 1989) namely:

$$\frac{dC_i}{dt} = \Phi_i(C_1, \ldots, C_n) + D_i \nabla^2 C_i,$$

with $i = 1, \ldots, n$, the index associated to the i-th morphogen having a local concentration $C_i$. Each term $\Phi_i$ describes how the i-th molecular species reacts to the other molecules. The last term in the right-hand is the diffusion term accounting for the spontaneous, random movement of molecules through space. However, the formalism needs to be extended by incorporating a changing boundary which now acts as a permeable membrane, also coupled to the reactions described by $\Phi_i$. These reactions will define the protocell metabolism. Since osmotic pressures are associated to differences in molecular concentrations, active mechanisms generating spatial heterogeneity are expected to create changing pressure fields. These instabilities can break the osmotic pressure symmetry along the membrane, and after division the reactions defining the metabolism must be able to trigger a new growth-division cycle (Macía & Solé, 2006b).

3. Chemoton-like protocells: the chemoton introduced by Gánti (1975) consists in three functionally dependent autocatalytic subsystems: the metabolic chemical network, the template polymerisation and the membrane subsystem enclosing them all (fig. 5c). The self-reproducing metabolic network transforms the external nutrients into chemoton’s internal material necessary for template replication and membrane growth. The correct functioning of the chemoton lies in the precise stoichiometric coupling of the three subunits, more precisely the co-ordination between the accumulation of molecules and the surface increase in order to achieve an equilibrium of the osmotic pressure relative to the environment. The model imposes a closed stoichiometric coupling of the autocatalytic cycles such that the number of membrane molecules necessary for
surface doubling is equal to the number of polymerisation iterations needed for a complete replication of all double-stranded template molecules. As shown by recent works (Table 1 – 17), the stoichiometric coupling is not necessary for fulfilling coordinated growth as the catalytic coupling of the subsystems can ensure their coordinated growth, and thus the doubling of their components prior to division, which is the condition for a viable replication cycle.

4. Ribo-cells: (fig. 5d) Szostak et al. (2001) and Bartel & Unrue (1999) have suggested possible RNA protocols under the form of minimal ribo-organism: one encapsulated ribozyme would synthesise the vesicle membrane component, and a second ribozyme would replicate itself and the first one. The components necessary for the RNA implementation as seen in these works are not yet available, and there are clues suggesting that a DNA cell might appear easier to implement experimentally than “simpler” RNA cell (Luisi et al., 2002). However, a different RNA cell has been proposed and implemented in the laboratory by Chen et al. (2004) showing that RNA replicating within vesicles could increase membrane growth rate by creating internal osmotic pressure. An even bigger step towards the prototype of the minimal protocol was accomplished by Ishikawa et al. (2004) through the laboratory implementation of a two-level cascading protein expression in liposomes. The simple structure of ribocells allows them to be fully described by means of whole-protocol simulation models (see for example Flamm et al. 2006).

5. Transduction-driven protocells: Efficient protocellular systems incorporating both DNA and RNA (fig. 5e) could be obtained by incorporating appropriate energy transduction systems (Pohorille et al., 1996). One such subsystem appears indicated within the protocol in fig. 5e. It consists of two proteins, bacteriorhodopsin (BR) and the FOF1 ATPase from the thermophile Bacillus PS3, embedded in liposomes (Richard et al., 1995). BR is a light-driven proton pump. It generates the transmembrane proton gradient required for the ATP-making activity of the ATPase. It has been shown to have high turnover and stability. This would be a minimal protocell incorporating compartments inside its structure, thus making metabolism to be effectively associated with a special protocell substructure.

6. Minimal cells: the concept of minimal cell (fig. 5f) is the target of the works employing the top-down approach, as mentioned in the beginning of the previous section. Even though it appears as a more promising and straight-forward strategy in the search of a minimal protocell, when compared to the bottom-up approaches, there is still a lot of experimental work needed in order to establish tenable and incontestable rules of thumb for the behaviour of genetic modules within the genetic regulatory networks. Latest advances in synthetic biology have revealed opposing facts: that there is an enormous potential for designing and building programmable genetic devices, on one hand, and that surprises are to be expected when passing from simple circuits to devising multimodal or hierarchical genetic networks as a consequence of emergent new behaviours of numerous interacting agents, on the other hand. Thus, the step from simple genetic circuits to a new artificial cellular entity has to deal with both the unlimited behavioural diversity and the precision in reaching the programmed target intended for the artificial cell.

Most of these basic cell models can be used to implement computational tasks. Here computation means some sort of predictable response to external signals. In figure 6 we show two simple examples of explicit designs of two basic logic gates (the NOT and NOR gates, respectively) from the first model described in this section. Here the external signals correspond to some type of inhibitors of enzyme activity. As a measure of the output, we use cell division: the output $D$ will be one if division takes place and zero otherwise. In other possible scenarios, the output can correspond to some type of produced molecule resulting from cell reactions.

The NOT gate can be obtained by using an enzyme having a single inhibitor $I$: under the presence of $I$ no division occurs, whereas if absent the vesicle experiences reproduction. A simple generalisation of this using two enzyme inhibitors (here indicated as $I_1$ and $I_2$) leads to a NOR gate: unless both are absent, no reproduction can occur. Although these are rather trivial examples, they illustrate possible ways of designing simple types of computational cell structures. Once we incorporate as inputs the produced molecules or alternatively introduce different types of cells responding to different signals, it is not difficult to generate (at least at this theoretical level) more complex systems able to describe switches or even memory structures. It is worth mentioning that very complex computational devices (including Turing machines) are currently being developed experimentally at the molecular level (Shapiro & Benenson, 2006). These molecular automata might benefit in the future from the current advances in synthetic protocell biology.

V. DISCUSSION

The previous models and theoretical approaches to protocellular systems define a range of complexity levels from non-evolving, non-replicating structures to fully reproducing and evolving entities. So far none of these systems has been shown to exhibit autonomous reproduction. Achieving such goal would represent a great leap
in biology: it would provide the logical basis for understanding the nature and requirements of life at its simplest level. All the evolutionary advantages of cellular life (from compartmentalisation to division of labour) would enhance the potential advantages of synthetic molecular systems.

It is important to mention that possible scenarios of vesicle change in nature are not restricted to whole cell changes associated to reproduction. A wide diversity of mechanisms of membrane formation and processing exist inside complex cells, representing a variety of possible forms of building vesicles. These include transport dynamics using Golgi vesicles, endosomes, lysosomes and clathrin-coated vesicles (Alberts et al., 2002). Inspiration from such processes might help designing new types of protocellular systems. Some of these processes include the participation of the ribosome, a complex nanomachine involved in reading RNA molecules and translating them into proteins. The computational power of ribosomes, combined with designed protocellular systems exploiting spontaneous pathways of vesicle dynamics might help expanding the current state of the art in this area. Additionally, vesicles obtained from non-biological polymers have been shown to allow interfacing with biological structures (Discher & Eisenberg, 2002). These *polymersomes* can mimic many biological processes, including encapsulation of relevant molecules and transfer-loading through membrane proteins. Their enormous versatility allow the potential exploration of a highly diverse universe of hybrid biomembrane designs.

The success of any of the previous model approaches (and perhaps others not considered here) will provide the basis for a new field at the crossroads between biology and computation. Travelling from non-living to living matter means to cross a twilight zone: some transition domain where the preconditions for reliable cell replication (and thus life) exist. Although some steps need to be completed and some key processes are not yet understood, we are likely to see the success of synthetic cellular life soon at work over the next decade.

**Acknowledgments**

The authors would like to thank the members of the Complex Systems Lab for useful discussions. Special thanks to Jurgen F. Sebastian for useful discussions. This work has been supported by grants FIS2004-0542, ISTFET PACE project of the European Community founded under EU contract FP6-0022035 (Programmable Artificial Cell Evolution) and by the Santa Fe Institute.
References


Serra, R., Carletti, T. & Poli, I. 2006 (In press), Synchronization phenomena in surface-reaction models of protocells, *Artificial Life*.


