STOCHASTIC SIMULATIONS OF MINIMAL CELL MODEL SYSTEMS

Mavelli Fabio*, Margherita Caputo*, Altamura Emiliano*, Stano Pasquale°

*Chemistry Department, University Aldo Moro, Via Orabona 4 – 70125 Bari
° Science Department, University Roma Tre, V.le Marconi 446 - 00146 Roma

ABSTRACT

The construction of artificial cells based on the encapsulation of chemical reacting systems inside lipid vesicles is rapidly progressing in recent years. Several groups are currently interested in synthesizing such simple cell models for biotechnological purposes or for investigating origin of life scenarios. Within this context, the properties of lipid vesicles (e.g., their stability, permeability, growth dynamics, potential to host reactions or undergo division processes…) play a central role, in combination with the dynamics of the encapsulated chemical or biochemical networks. Thus, from a theoretical standpoint, it is very important to develop deterministic equations in order to explore first - and specify later - the conditions that allow the robust implementation of these complex chemically reacting systems, as well as their controlled reproduction. Due to their intrinsic compartmentalized nature, the population of reacting molecules can be very low in terms of number of molecules so that their behaviour can be highly affected by stochastic effects both in the time course of their reactions and in their occupancy distribution among the vesicle population. In this contribution we report our mathematical approaches to model artificial cell systems in this complex scenario, with emphasis on the issue of primitive cell (protocell) systems.

INTRODUCTION

The chemical implementation of diverse proto-cellular model systems is gathering the interest of a growing number of researchers in the fields of synthetic biology and origins of life[1]-[6], who are becoming aware of the potential of micro-compartments and lipid vesicle technologies to uncover biologically relevant phenomena, as well as prebiotically plausible processes and evolutionary transitions. Protocells are lipid micro-compartments (generally lipid vesicles, but other compartments have been also used) which contain a minimal number of (bio)chemicals in order to generate typical cellular behavior, like self-maintenance and self-reproduction.

Lipid vesicles consist in a closed, spherical, semi-permeable membrane formed by the spontaneous self-assembly of lipid molecules. The membrane is a highly organized molecular bilayer that separates the molecules trapped inside the vesicle (i.e., in the inner aqueous vesicle core) from the environment.

A variety of chemical and biochemical reactions have been implemented inside protocells, from RNA synthesis to gene expression, from DNA amplification to lipid synthesis (for a review, see [7]). The latter reaction is particularly important because it allows the growth of vesicles thanks to the enlargement of vesicle membrane. Division might also follow vesicle growth, so that two 'daughter' vesicles are obtained from a parent one (i.e., self-reproduction).

The aim of this contribution is to introduce the mathematical framework used to describe the time behaviour of reacting protocells in terms of the deterministic versus the stochastic approach [8] and to review and discuss some recent results obtained by our research group, focusing on the interplay between internalized reactions, vesicle growth and self-reproduction.

MATHEMATICAL BACKGROUND

In Silico Protocell Model

According to the schematic draw of a lipid vesicle reported in Fig. 1, reacting vesicles are described as compartmentalized systems made of two different homogeneous domains: the membrane and the water core [9]. Lipids can be exchanged between the membrane and water core and between the membrane and the external environment while transport processes can also occur, exchanging molecules directly from the external environment to the internal water pool. The vesicle membrane surface S_μ can be determined by its composition:

$$S_\mu = \sum_i^{Lipids} \frac{\alpha_i n_{\mu,1}}{2} \tag{1}$$

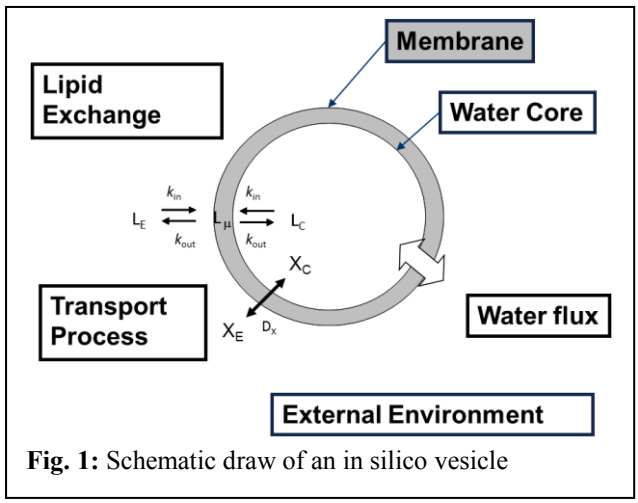

Fig. 1: Schematic draw of an in silico vesicle

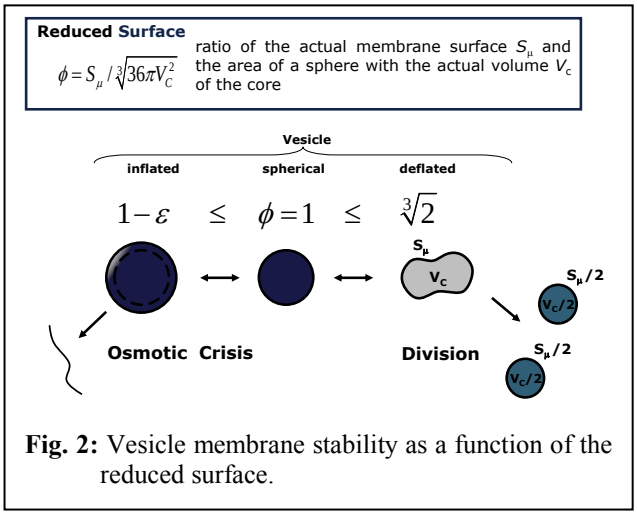

Fig. 2: Vesicle membrane stability as a function of the reduced surface.

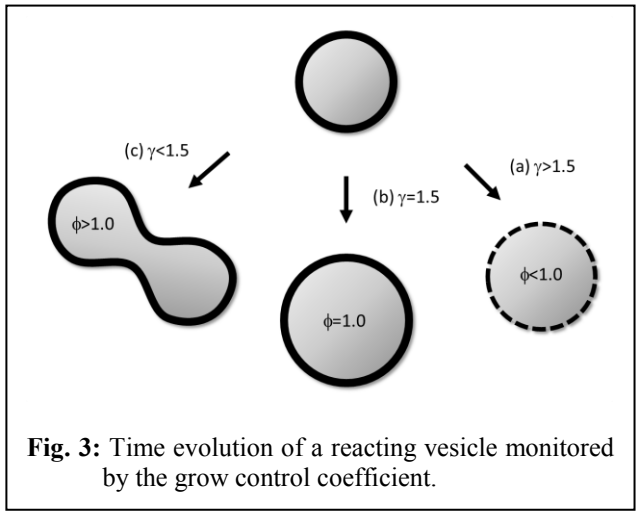

Fig. 3: Time evolution of a reacting vesicle monitored by the grow control coefficient.

according to the hydrophilic head area α_i of different lipids, while the internal aqueous volume core V_C is affected by a water flux due to osmotic pressure unbalance. In the rest of this paper we will deal with a vesicle membrane made of one lipid molecule so the previous equation simplifies $S_\mu = \alpha_L n_L^\mu / 2$. The membrane stability can be monitored by introducing the reduced surface ratio ϕ, that equals 1.0 for spherical vesicles, while it will be less than 1.0 or greater than 1.0 for inflated or deflated vesicles respectively. In fact, a flux of water can takes place across the lipid membrane driven by an osmotic pressure unbalance. Therefore, inflated vesicles can undergo an osmotic crisis when the internal volume grows to much bringing the membrane to rupture when $\phi<(1-\varepsilon)$, ε being the osmotic tolerance. On the other hand, in the present model deflated vesicles are assumed to divide when the membrane surface is large enough to form two twin spherical daughters: $\phi = \sqrt[3]{2}$. This event has been observed in some experimental conditions [10] nevertheless the dynamics of a deflated membrane is a much more complex process.

Since the aqueous core volume V_C and the membrane surface S_μ may follow independent time trends, in order to describe the various possible behaviors of the system, it is convenient to introduce the *growth control coefficient* γ [11]:

$$\gamma = \left(\frac{1}{V_C}\frac{dV}{dt}\right) \Big/ \left(\frac{1}{S_\mu}\frac{dS}{dt}\right) = \frac{S_\mu}{V_C}\frac{dV}{dS} \tag{2}$$

This dimensionless observable is defined as the ratio between the relative velocities of variation of volume and surface, respectively: In presence of an endogenous (biosynthetic or proto-metabolic) production of lipid results $dS>0$ thanks to the spontaneous uptake of fresh lipid by the membrane and, in these conditions, $\gamma>0$ indicates a real growth regime. Therefore, just by applying some straight-forward geometry rules for a growing sphere: $d(lnV)/3 = d(lnS)/2 = d(lnR)$, three different scenarios among all possible growth regimes may be distinguished: (a) $\gamma=3/2$ *continuous spherical growth*, i.e. a spherical vesicle will increase its size without any change of shape ($\phi=1$); (b) $\gamma>3/2$ *osmotically-stressed growth*, i.e. the volume increases faster then it will reach an elastic tension condition and, above the limit of elasticity of the membrane, this will lead the vesicle to osmotic burst($\phi<1-\varepsilon$); (c) $\gamma<3/2$ *reproductive growth*, i.e. the surface increases faster than the two previous cases, the growing vesicle will turn deflated, changing to some other closed but non-spherical shape (ellipsoidal, elongated or, generally speaking, a prolate shape) and the energy of the membrane will be higher due to a bending tension.

Deterministic Approach

If in the internal core of the compartment, N species X_i (i=1,2…N) react according to R chemical elementary reactions:

$$r_{1,\rho}X_1 + r_{2,\rho}X_2 + \ldots + r_{N,\rho}X_N \xrightarrow{k_\rho} p_{1,\rho}X_1 + p_{2,\rho}X_2 + \ldots + p_{N,\rho}X_N \qquad \rho=1,2..R$$

then the average time evolution of the reacting vesicles can be described by the deterministic approach [8] solving the following ordinary differential equation set (ODES):

$$\begin{cases} \dfrac{dx_i}{dt} = N_A V_C \sum_{\rho=1}^{R}\left(p_{i,\rho} - r_{i,\rho}\right)v_\rho + N_A \wp_i S_\mu \left(\left[X_i\right]_{Ex} - \dfrac{x_i}{N_A V_C}\right) \\ \qquad\qquad i = 1,2\ldots N, \\ \qquad\qquad i \neq L \\ \dfrac{dx_L}{dt} = N_A V_C \sum_{\rho=1}^{R}\left(p_{i,\rho} - r_{i,\rho}\right)v_\rho - k_{in} S_\mu \dfrac{x_L}{N_A V_C} + k_{out} x_L^\mu \\ \dfrac{dx_L^\mu}{dt} = k_{in} S_\mu \dfrac{x_L}{N_A V_C} - k_{out} x_L^\mu \\ \dfrac{dV_C}{dt} = \omega_{aq}\wp_{aq} S_\mu \left(C_C - C_{Ex}\right) \end{cases} \tag{3}$$

where v_ρ are the reaction rates given by the mass action law:

$$v_\rho = k_\rho \prod_{n=1}^{N}\left(\frac{x_n}{V_C N_A}\right)^{r_{n,\rho}} \tag{4}$$

The solution of the ODES gives the average time behavior of the vesicle solution in terms of the number of molecules x_i of internal aqueous species X_i (i=1,2…N, $i\neq L$), the lipid molecules x_L in the water core, the lipid molecules in membrane x_L^μ and core volume V_C. Moreover, it has been written for the case of aggregates formed by a single lipid X_L. Others simplifying assumptions are to neglect the specie diffusion in the internal core and in the external environment as well, and assuming the external concentration $[X_i]_{Ex}$ to be

constant in time, i.e. the environment is considered as an infinity source of external compounds.

Going into details, the mole number rate change of each aqueous species dx_i/dt is due to the internal metabolic reaction and to the transport process from the outside. The transport across the membrane is driven by a concentration gradient as shown by the following scheme:

$$X_i^{Ex} \xleftrightarrow{N_A \wp_i S_\mu \left(\left[X_i \right]_{Ex} - \left[X_i \right] \right)} X_i \qquad (5)$$

where $\wp_i$ is the membrane permeability of *i*-species and N_A is Avogadro's number. Instead, the rate change of the lipid in the core dx_L/dt takes into account the exchange between the aqueous internal phase and the membrane described as follows:

$$X_L \underset{k_{out} n_L^\mu}{\overset{k_{in} S_\mu \left[X_L \right]}{\rightleftarrows}} X_L^\mu \qquad (6)$$

while the lipid exchange towards the outside is not explicitly considered in rate change of membrane lipids dx_L^μ/dt since the external lipid concentration is assumed to be constantly equal to the equilibrium value $[X_i]_{Ex} = \alpha_L k_{out}/(2k_{in})$. The last equation in the ODES (3) describes the core volume rate change due to a flux of water driven by the difference of the total osmolite concentration, i.e. an osmotic pressure unbalance, being $\wp_{aq}$ the water permeability and ω_{aq} the water molecular volume. It is important to remark that the deterministic approach gives the time evolution of the vesicle solution as the average time course calculated over the vesicle ensemble, so that x_i and x_L^μ are not positive integer numbers but they are positive real values nevertheless; they represent amount of molecules, Therefore the vesicle state is represent by the array $\mathbf{X}=(x_1, x_2, \ldots, x_N)^T$ and the core volume V_C. When the condition for division is satisfied ($\phi = \sqrt[3]{2}$), then the vesicle divides in two twin daughters with volume equal to $V_C/2$ and all the elements of the state array are accordingly divided by 2.

Stochastic Simulations

The stochastic kinetic approach explicitly takes into account the discrete nature of molecules and the intrinsic randomness of reacting events. Therefore, the state of a reacting vesicle is defined by an array of integer molecular numbers n_i: $\mathbf{N}=(n_1, n_2, \ldots n_N, n_L^\mu)^T$ and the core volume V_C. Moreover, for each elementary reacting event a propensity density probability $a_\rho(\mathbf{N})$ is introduced instead of the deterministic reaction rate so that $a_\rho(\mathbf{N})dt$ gives the probability ρ-th reaction will take place in the next infinitesimal time interval dt [8]*:*

$$a_\rho(\mathbf{N}) = \frac{k_\rho}{\left(N_A V_C\right)^{\left(\sum_j^N r_{j,\rho}\right)-1}} \prod_j^N \binom{n_j}{r_{j,\rho}} \qquad (7)$$

while the propensity density probabilities for transport processes and lipid exchange can be predicted according to eqs. (5) and (6) [12]. The stochastic time evolution of a well stirred chemically reacting system can be then obtained by solving the Master Equation (ME) [8]:

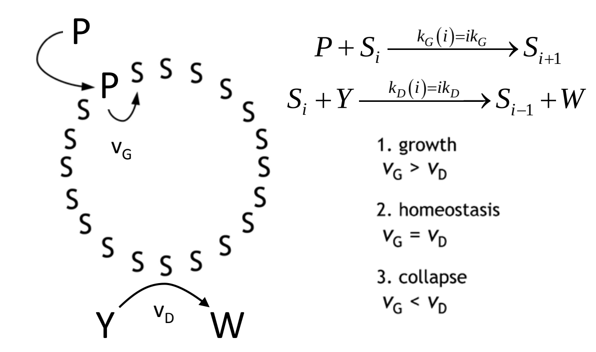

Fig. 4: Autopoietic vesicles: schematic draw, kinetic mechanism and different regimes.

$$\frac{\partial P\left(\mathbf{N},\, t \mid \mathbf{N}_0,\, t_0\right)}{\partial t} = \sum_{\rho=1}^{R} a_\rho(\mathbf{N} - \Delta\mathbf{N}_\rho) P\left(\mathbf{N} - \Delta\mathbf{N}_\rho,\, t \mid \mathbf{N}_0,\, t_0\right) + - P\left(\mathbf{N},\, t \mid \mathbf{N}_0,\, t_0\right) \sum_{\rho=1}^{R} a_\rho(\mathbf{N}) \qquad (8)$$

that expresses the change rate of the Markov Density function $P(\mathbf{N}, t| \mathbf{N_0}, t_0)$, i.e. the density probability to find the system in the state $\mathbf{N}$ in the time interval $[t, t+dt)$ given the system in the state $\mathbf{N_0}$ at time t_0. $\Delta\mathbf{N}_\rho$ is the jump array, that is the stoichiometric variation of the number of molecules due to the *ρ-th* reaction. By solving analytically the ME, the average time behavior of the reacting system can be obtained along with displacements from the average species time trend due to random fluctuations that can bring the system towards regimes unpredictable by the deterministic approach[8]. ME is very difficult to solve analytically, but it can be exactly simulated by the well know Monte Carlo direct methods introduced by Gillespie [16]. Based on this method we developed a software platform [12] suitable to simulate the stochastic time evolution of a collections of reacting vesicles assuming that diffusion processes can be neglected and the concentration gradients take place only across the lipid membrane. This program allows also to study the case of vesicle self-reproduction since it is able to follow a collection reacting compartment that increases in number. For further details the reader is address to references[12]-[15]. What we want to remark here is this program is suitable to study also the influence of extrinsic stocasticity. In fact, reacting molecules can be distributed randomly among compartments at the starting time or between daughters at the division time ($\phi = \sqrt[3]{2}$), simulating how this source of randomness affects the system time behavior.

RESULTS AND DISCUSSION

Autopoietic vesicles in homeostatic regime

Autopoiesis, as developed by Maturana and Varela in the seventies [17], is a theoretical description of the 'blue print' of cellular life. It poses as a main feature the self-maintenance of the cell, as due to a process of components' self-generation from within the cellular boundary—a boundary which is itself

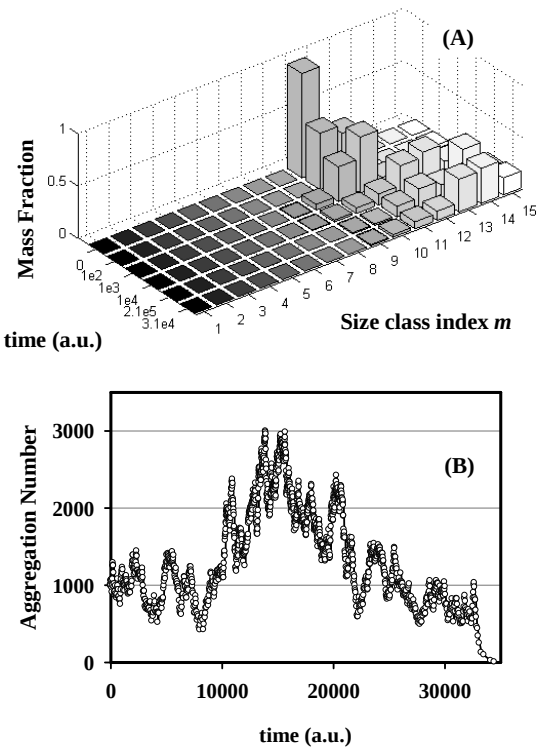

Fig. 5: Stochastic simulations of autopoietic vesicles in homeostatic conditions: (A) evolution of the vesicle size distribution (to each size class belong ennamers with size $2^{m-1} < i \leq 2^m$ except for the first class m=1 where only monomers are included); (B) time evolution of a single aggregate.

one of the products. From the chemical point of view, the fertility of autopoiesis theory allowed the design and the experimental achievement of some autopoietic chemical systems all based on surfactant self-assembling structures, such as micelles, reverse micelles and vesicles [18]. In Fig. 4, the schematic representation of an autopoietic vesicle is shown along with the kinetic conditions for experimentally observing different regimes in the time course of total surfactant concentration depending on the rates of amphiphiles production v_G and decay v_D respectively. These three scenarios have been really implemented and investigated by Zepik et al. [19]. In particular, the chemical system consists in a solution of oleic acid/oleate vesicles (S_i, i being the aggregation number), buffered at pH 8.8, fed with a surfactant precursor and with a reactant capable of destroying oleic acid. The surfactant precursor (P) is oleic anhydride, a hydrophobic substrate rapidly taken up by oleate vesicles at their membranous interface. Thanks to the high pH value, P is converted to oleate by alkaline hydrolysis that takes place on the membrane of vesicles. Oleate vesicles also undergo a decay process due to the simultaneous transformation of oleate molecules into 9, 10-dihydroxystearate (W) by osmium tetroxide/potassium ferrocyanide oxidation (Y). The dihydroxylated compound P does not form vesicles; therefore, the consequence of the latter conversion is a stepwise vesicle collapse (death). Due to the two competitive reactions, the overall oleate concentration increases, remains approximately constant, or decreases, depending on the magnitude of the P and Y flux rates [19]. In order to reproduce the experimental

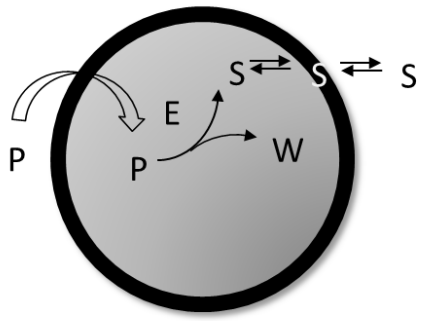

Fig. 6: Self-Replicating Enzymatic Vesicle: P lipid precursor, S surfactant, E enzyme and W waste.

observed behavior we proposed the simple mechanism reported on the right of Fig. 4 and we was able to obtain the time course of the overall oleic acid concentration [15]:

$$[S] = [S]_0 \, e^{\left(k_G[P]_0 - k_D[Y]_0\right)t} \tag{9}$$

This equation accounts for the three regimes by explicitly expressing the rates of amphiphiles production $v_G = k_G[P]_0$ and decay $v_D = k_D[Y]_0$, as a function of the aqueous concentration of the anhydride $[P]_0$ and of the oxidant $[Y]_0$ kept constant by the external fluxes. Stochastic simulations performed in homeostatic conditions ($k_G[P]_0 = k_D[Y]_0$) have been then done in order to elucidate the evolution of the vesicle size distribution.

Starting from a size monodispersed ennamers solution, what emerges from simulations is that stochasticity selects ennamers with aggregation numbers in the range 10^3–10^4 (Fig. 5A). This effect can be ascribed to the presence of random fluctuations in the growth and decay specific rates, which in real (chemical) reacting systems are due to the intrinsic stochasticity of reacting events but they can also be enlarged by natural changes of physical parameters such as temperature, molecular fluxes, etc. In fact, stochastic simulations starting from a single aggregate have shown how random fluctuations at the steady state can drive the evolution of the aggregate towards a growth or a decrease in size (Fig. 5B). Therefore, when autopoietic ennamers of different sizes are present in a system in stationary conditions, fluctuations can act as a selection rule that leads to the perpetuation of those aggregates large enough to overcome large deviations.

In conclusion, stochastic simulations have shown that, in this landscape, random and driven fluctuations can represent the driving force for ennamer evolution, growth or decay, and at the same time they can act as a selection rule for the fittest, i.e. the most robust, aggregates in a prebiotic environment.

Protocell stationary self-reproduction

In a recent work, a phenomenological law that predicts when a *stationary self-reproduction* takes place for minimal self-producing vesicles have been derived. By 'stationary self-reproduction' we mean a dynamic regime where the condition for division is reached at a constant, characteristic period of time, giving as a result two vesicles or *protocells* with the same (initial) size, lifetime and metabolite concentration profile as the progenitor.

In terms of the *growth control coefficien*t the steady condition takes place when γ=1. Then, two general

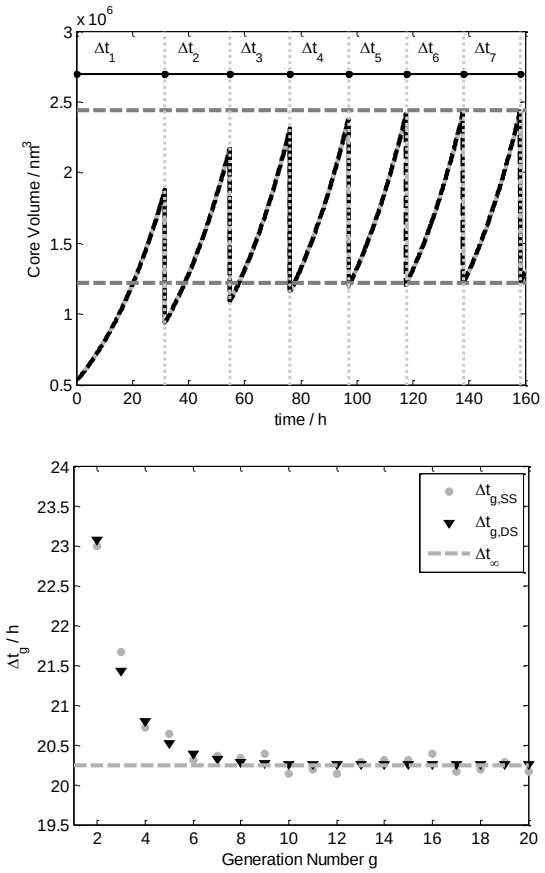

Fig. 7: Self-Replicating Enzymatic Vesicle deterministic curves (black lines and data) and stochastic simulation results (gray lines and data) comparison: time evolution of the core volume (top plot); division time against generations (bottom plot). Horizontal dashed lines represent values calculated by eqs.(12) and (13).

expressions for the temporal behavior of the protocell surface and the protocell core volume have been independently derived [11] and an explicit relationship among different molecular and kinetic parameters (e.g., reaction rates v_ρ, permeability coefficients $\wp_i$, metabolite concentrations $[X]_i$) have been analytically derived for the protocell stationary reproduction:

$$\sum_{\rho}^{\text{Reactions}} \Delta m_\rho v_\rho + \frac{S_\mu}{V_C} \sum_{i}^{\text{Species}} \wp_i \left([X_i]_{Env} - [X_i] \right) - v_L = \frac{C_C \alpha_L}{2} \frac{N_A V_C}{S_\mu} v_L \tag{10}$$

where v_L is the rate of lipid production, C_C is the total internal concentration and Δm_ρ is total variation of the number of molecules due to the ρ-*th* reaction:

$$\Delta m_\rho = \sum_i (p_{i,\rho} - r_{i,\rho}) \tag{11}$$

Eq.(10) shows the deterministic condition for a stationary reproduction regime that results from the *osmotic synchronization* between membrane and core volume growth, i.e.: a spontaneous 'self-regulation' driven by the osmotic balance across the protocell lipid bilayer. Eq.(10) links metabolic kinetic constants and membrane permeabilities with the external and internal concentrations of the system constituents. Therefore, it represents a constraint for the possible sizes and division periods of stationary self-reproducing protocells. We have applied the general eq. (10) to the simplest case of a self-producing enzymatic vesicle (SPEV) represented in Fig. 6. SPEV is a hypothetical protocell model where the production of lipid S takes place through the chemical transformation of a precursor molecule P, assumed to occur only in the presence of an additional compound E encapsulated in the core volume. The S production generates also the waste W so that Δm=1 and the osmotic synchronization can in principle takes place. Moreover W is accumulated in the core volume since it is assumed not to be transported across the membrane, i.e. $\wp_W$=0. It is worthwhile to note that this model is very close to some experimental approaches based on giant vesicles that produce internally (with the help of a synthetic catalyst) the main membrane component and eventually undergo self-reproduction [20].

SPEV is not a real autopoietic vesicle since the catalytic specie E is not synthetizes by the internal metabolism. Therefore after each vesicle division the number of E molecules will decrease until just one copy of these molecules will be present in the internal core. As a consequence, whenever a division occurs only one of the two daughter vesicles will be able to encapsulate the catalyst molecule and, therefore, will keep the potential to continue growing, producing S and, eventually, reproducing as a protocell. The vesicle that contains that single molecule E, by default, will be taken as the mother vesicle, whereas the daughter (and all possible granddaughters) will be 'sterile' vesicles. Thus, by handling eq. (10) it was possible to predict [11] for the mother SPEV, i.e. the vesicle containing only one E molecule, its stationary radius R_∞:

$$R_\infty = \frac{6}{\alpha_S N_A C_C} \tag{12}$$

and the division time Δt_∞:

$$\Delta t_\infty = \frac{2}{\alpha_S [P]_{Ex}} \left(\frac{\ln 2}{N_A \wp_P} + \frac{4\pi R_\infty^2}{k} \right) \tag{13}$$

where C_C is the overall internal osmolite concentration, $[P]_{Ex}$ and $\wp_X$ are the external concentration and the membrane permeability of the lipid precursor respectively, while k is the kinetic constant of the lipid production: $v_L = k[E][P] = k[P]/(N_A V_C)$.

Fig. 7 shows in the upper plot the core volume time trend for the first 7 generations, i.e. vesicle divisions, obtained both by ODES integration (black line) and by stochastic simulations (gray lines). Vertical gray dotted lines represent the time of division that takes place when the reduced surface satisfied the splitting conditions: $\phi = \sqrt[3]{2}$. Generation by generation the mother protocell tends to the stationary growth and division as illustrated by the upper plot where the core volume values before $2V_\infty$ end after the division $V_\infty = 4\pi R_\infty^3/3$ can be calculated with eq.(12). In the lower plot it is reported the division time Δt_g against the generation number, showing that generation by generation it tends to Δt_∞ as predicted theoretically.

An important aspect to remark is that eq.(10), strictly speaking, only captures the condition for stationary reproduction in the sense of a global synchronization process between membrane and volume growth. In other words, it does not guarantee that when a vesicle reaches the division threshold the number of each internal constituent gets effectively doubled (with regard to their initial state in the

protocell cycle). This becomes manifest in the case of SPEV, where the single enzyme/catalyst present in the mother is not doubled and, therefore cannot be transferred but to one of the offspring vesicles (i.e.: the only one that will remain fertile). Therefore, eq.(10) states a necessary but not sufficient condition for reliable reproduction of proto-cellular systems. In a more complex scenario, which will be introduced in the next section and where the metabolic reaction network included the synthesis of the enzymatic/catalytic compound, a more complete reproduction of the protocell could be achieved. But the synchronization among lipid production, enzyme duplication and membrane division would emerge in that system only if the new metabolic pathway(s) lead to effective internal chemical synthesis (i.e. $\Delta m>0$), since the mechanism that drives the synchronization is the osmotic balance across the lipid bilayer. Moreover in this complex scenario also stochastic fluctuations can effect much more the time behavior of each single protocell and in particular the random distribution of enzymatic species between daughter vesicles.

A minimal cell model: the Ribocell

The so-called Ribocell (RNA-based cell) is a theoretical minimal cell model based on a self-replicating minimum RNA genome coupled with a self-reproducing lipid vesicle compartment that has been recently hypothesized [21]. This model suppose the existence of two ribozymes, one (the lipid synthase R_L) able to catalyse the conversion of molecular precursors (P) into lipids (S) and the second (the polymerase R_P) able to replicate RNA strands by a template driven elongation. Therefore, in an environment rich in both lipid precursors (P) and activated nucleotides (NTP), the Ribocell can self-reproduce if both processes, i.e. genome self-replication and membrane reproduction (growth and division), are somehow synchronized. Recently we have explored the feasibility of this hypothetical minimal cell [14] by determining the best external conditions to observe synchronization between genome self-replication and vesicle membrane reproduction, thanks to a deterministic kinetic analysis, while the Ribocell robustness to random fluctuations has been tested by stochastic simulations. The proposed metabolic mechanism is reported in Fig. 8. Both pairs of RNA strands reversibly associate (A) and these processes are shifted towards the dimer formation and are strongly dependent on temperature. The replication of any RNA strand is catalysed by the polymerase R_P according to the steps in bracket (B). This process is described as a catalytic template-directed addition of mononucleotides with high fidelity and processivity. It starts with R_P binding any of the monomeric template T (T=R_P, $_cR_P$, R_L and $_cR_L$) to form the complex R@T. This complex will then initiate the polymerization of the conjugate strand $_cT$, by coupling and iteratively binding the complementary bases and releasing the by-product W. When the strand $_cT$ has been completely formed, the polymerase ribozyme releases the new dimer. Finally, the ribozyme R_L catalyzes the conversion of the precursor P into the lipid S (C). All the kinetic constants have been estimated by experimental values reported in literature and are listed in Table 1 along with references.

Thanks to a deterministic analysis [28][29], we showed that if the kinetic constant for lipid formation k_L is in the range: $1.7\cdot10^3 s^{-1}M^{-1} \leq k_L \leq 1.7\cdot10^5 s^{-1}M^{-1}$ then synchronization between vesicle reproduction and genome replication can

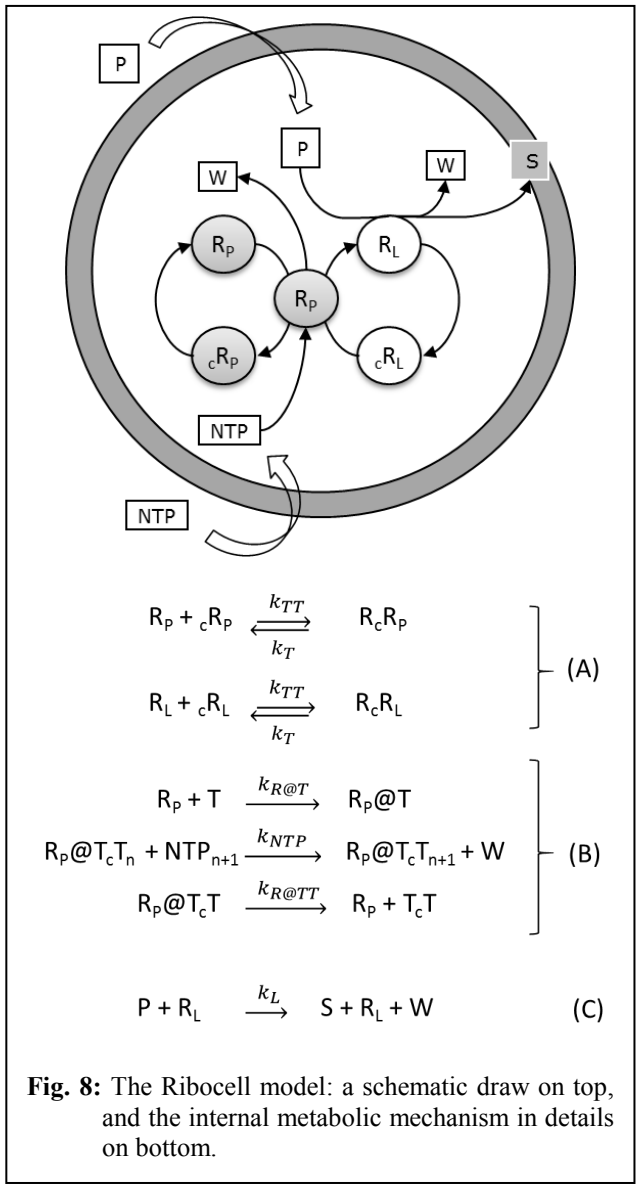

Fig. 8: The Ribocell model: a schematic draw on top, and the internal metabolic mechanism in details on bottom.

spontaneously emerge under the model assumptions and kinetic parameters reported in Table 1. Deterministic calculations were performed for two ribozymes 20 bases long and showed that the Ribocell reaches a stationary growth and division regime (γ=1), where the cell size remains constant after each division along with the amount of genetic material. Although the observed cell life time stabilizes after the first 10 generations, it remains very high, at over 80 days for all the k_L values in the synchronization range, making the Ribocell very hard to implement and study experimentally. Therefore, we investigated the robustness of the stationary growth and division regime of the Ribocell in terms of the external substrate concentrations, vesicle size and initial ribozyme amount in order to define optimal external conditions for Ribocell self-reproduction [14]. The influence of ribozyme length will also be explored in the optimal external conditions by ranging strand size from 20 to 200 bases in length and keeping all the other kinetic parameters constant. 20 bases is in fact the minimum length required to observe a folded RNA structures, i.e. a structure that can reasonably exhibit catalytic action. On the other hand, entities of about 200 nucleotides have been suggested as plausible ancient proto-ribosomes [30] even though, more recently, smaller subunits of 60 nucleotides have also been considered as plausible candidates [31]. This analysis shows that starting from external concentrations

Table 1: Kinetic Parameters for the in silico Ribocell model at room temperature.

Kinetic Parameters	Values	Ref.
$K_{TT}[s^{-1}M^{-1}]$	$8.8 \cdot 10^{6}$	[22]
$K_T[s^{-1}]$	$2.2 \cdot 10^{-6}$	[22]
$k_{R@T}[s^{-1}M^{-1}]$	$5.32 \cdot 10^{5}$	[23]
$k_{R@TT}[s^{-1}]$	$9.9 \cdot 10^{-3}$	[23]
$k_{NTP}[s^{-1}M^{-1}]$	0.113	[25][26]
k_L $[s^{-1}M^{-1}]$	$1.7 \cdot 10^{3}$	[24]
k_{in} $[dm^2s^{-1}]$	$7.6 \cdot 10^{19}$	[12]
k_{out} $[dm^2s^{-1}]$	$7.6 \cdot 10^{-2}$	[12]
$\wp_P$ $[cm \cdot s^{-1}]$	$4.2\ 10^{-9}$	
$\wp_{NTP}$ $[cm \cdot s^{-1}]$	$1.9\ 10^{-11}$	[25][26]
$\wp_W = \wp_T$	0.0	
$\wp_{aq}[cm \cdot s^{-1}]$	$1.0 \cdot 10^{-3}$	[27]

* k_L is 10^5 times larger than the value of the splicing reaction, catalyzed by the hammerhead ribozyme.

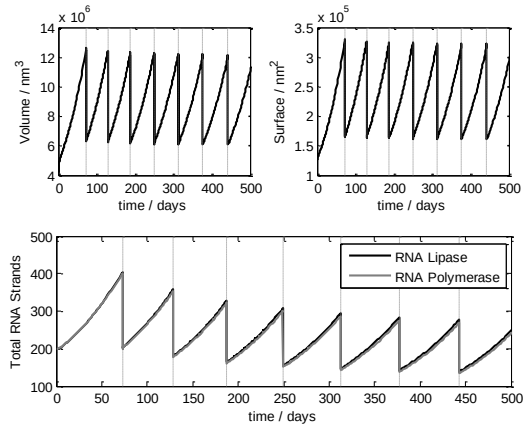

Fig. 9: Deterministic time behavior of the Ribocell in optimal external conditions: $[NTP]_{Ex}=[P]_{Ex}=0.01M$. At the starting time the genome was composed by 100 dimers of R_cR_L and R_cR_P and the radius was 100nm and the core volume $4.2\times10^6 nm^3$.

$[NTP]_{Ex}=[P]_{Ex}=10^{-2}M$ at the stationary regime the Ribocell radius is 113.0 nm and the division time reduces to 68.2 days. The total number of RNA strands is 258 and the genome composition is quite uniform 25.2% (R_L), 25.2% ($_cR_L$) 25.6% (R_P), 24.0% ($_cR_P$). The stationary division regime can be reached starting from initial genome composition ranging from 1 to 100 dimers of R_cR_L and R_cR_P. In Fig. 9, the deterministic time behavior of the Ribocell in optimal external conditions is reported.

Finally, the dependence of Δt_{25} (division time after 25 generations) on the kinetic constants for RNA dimer formation k_{TT} and dissociation k_T has been also studied. What emerged is the Ribocell life cycle at stationary regimes does not depend explicitly on the kinetic constant single values k_{SS} and k_S but on their ratio: k_{SS}/k_S, that is on the thermodynamic constant of RNA dimerization. The more thermodynamically stable the RNA dimers, the longer it takes to observe Ribocell self-reproduction. For instance, if k_{SS}/k_S is decreased by two orders of magnitude, the Ribocell life time reduces from 68.2 days to 11.8-6.4 days.

Stochastic simulations have been then performed in order to test the robustness of the ribocell base on 100-base length ribozymes in optimal external conditions, with the aim of elucidating the role of intrinsic and extrinsic stochasticity on the time behavior of a protocell population. Simulations were executed by means of the parallel version of ENVIRONMENT [12], running 32 statistically equivalent simulations of a 10-ribocell solution on different CPUs. Therefore, the outcomes were obtained as averages from a population of 320 vesicles. Kinetic parameters used for simulations are those reported in Table 1. At each cell division, only one of the two offspring was kept while the other was discarded in order to reduce computation time, thus keeping the number of monitored vesicles constant. This is in agreement with the assumption that the external concentrations of all substrates are fixed due to an incoming flux of material, i.e. the substrates cannot ever be exhausted. The simulation outcomes are reported on the left of Fig. 10 where the composition of the Ribocell population is reported against time. In fact, during simulations at each division the genetic material is randomly distributed between the daughters. If the amount of genetic material is very low, then this can result in a separation of R_P from the other RNA strands. In fact, the Ribocell must contain a minimum genetic kit of three RNA filaments in order to be capable of self-replicating its entire genome: one R_P that catalyzes the RNA base pair transcription, one (R_L or $_cR_L$) and one (R_P or $_cR_P$) that work as templates for the transcription. Moreover, since R_L is necessary to catalyze lipid precursor conversion, the optimal minimum 3-ribozyme kit must be made up of $2R_P$ and one R_L. This minimum kit should be at least doubled before cell division, in order to have a chance that both daughters continue to be active. Therefore, if a random distribution of RNA filaments takes place after vesicle division, ribozyme segregation between the two daughters might occur. Different scenarios can be envisaged: *death by segregation* is reached if vesicles are produced without any ribozymes (*empty vesicles*) or containing one lone R_P or many filaments of $_cR_P$ and/or $_cR_L$ (*inert vesicles*). Vesicles that encapsulate R_L strands are *self-producing*: they are able to synthesize lipids and then can grow and divide producing in turn self-producing and/or empty vesicles. On the other hand, vesicles containing more than one molecule of R_P or both R_P and $_cR_P$ filaments are able to self-replicate this reduced genome (*self-replicating genome vesicles*) but they cannot self-reproduce the membrane. So they are destined for an osmotic burst due to an unbalanced increase in waste concentration. Finally, a reduced version of the Ribocell consists in a lipid aggregate that contains one R_P filament and $R_L/_cR_L$ strands. As a consequence of this, *reduced ribocells* are able to replicate the $R_L/_cR_L$ genetic stuff, and at the same time to synthesize lipids. Therefore, they can grow and divide, producing in turn at least one reduced ribocell and/or self-replicating, inert and empty vesicle.

On the left of Fig. 10 a schematic draw of the different types of protocells is reported. At the end of the simulation, the composition of the protocell population are obtained with low percentages of real ribocells (6.7%) while the most populated fractions are those of empty (40.0%) self-producing (33.3%) and broken (20.0%) vesicles, respectively. Reduced ribocells are present only in the first generations since they very soon decay into self-producing and empty vesicles. Inert vesicles, i.e. vesicles entrapping free chains of cRP and/or cRL or a single RP, are not formed and this can be ascribed to the high stability of RNA dimers and complexes so that the chance of finding free RNA monomers at the time of vesicle division is extremely improbable. Indeed, the stochastic time

trend presents a very irregular time behaviour compared to the deterministic one that describes a highly synchronized oscillating regime of growth and division. In contrast, stochastic simulations highlight the alternation of dormant phases, where the reduced surface remains practically constant, both the core volume and the membrane surface being constant (data not shown), to very active steps where protocell growth takes place very fast, leading to a division event. The fast growth and division step corresponds to the presence in the vesicle core of a free R_L chain while, in the dormant phase, ribozymes are all coupled in the form of dimers or complexes. As a consequence, self-producing vesicles with a genome made up only of R_L monomers can reproduce very efficiently since no dormant phase can occur, given that the formation of R_cR_L dimers is impossible. This protocells could then self-produce very efficiently, with a Δt less than one day.

In conclusion, the simulation outcomes show that ribocells are not enough robust to survive to random fluctuations. In fact only about the 5-7% of the initial population survive as genuine ribocells after 15-25 generations and on a longer time window they are destined for extinction. Furthermore, the time course of each single protocell is also greatly influenced by intrinsic stochasticity in particular by the time fluctuations of the RNA dimer dissociation. In fact, when all the RNA strands are associated in dimers, protocells remain in a lazy phase, whereas free R_L monomers induce fast growth and division steps and free R_P cause the fast RNA replication without changing the vesicle size appreciably. Therefore these two processes are synchronized only by chance and this also represents a reason of weakness of this model protocell. Further details can be found in papers [14], [28] and [29]

CONCLUSIONS

In this short article we have shown some aspects of theoretical modelling in micro-compartmentalized systems, and in particular in the research on self-reproducing protocells. The occurrence of compartmentalized synthetic reactions coupled with the membrane dynamics in terms of growth and division plays a major role in determining the evolution of the system. In particular, we have firstly compared the deterministic and stochastic approach for modelling such systems, and applied these methodologies to describe (1) homeostatic autopoietic systems, (2) the stationary conditions for protocell self-reproducion, and (3) the more complex case of the "ribocell", i.e., a protocell based on catalytic function encoded in self-reproducing ribozymes (RNA enzymes).

Here we would like to emphasize the common aspects of analysis and modeling of these (and other) systems, namely the need of a systemic approach that integrates (and couple) the internal reactions, the membrane dynamics, and the environment. This is perhaps the most important scientific message that emerges from numerical simulations of these complex systems. Since numerical modeling is carried out by using true physical constants for all elementary molecular steps, it follows that genuine outcomes from modeling might actually help the experimentalists to design and construct protocell models or artificial cells for nanotechnological applications. Moreover, flanking stochastic modeling to deterministic approaches uniquely reveals intriguing dynamics in microcompartmentalized complex multimolecular systems and greatly helps to evaluate and understand basic mechanisms at the roots of biological behaviour.

ACKNOWLEDGMENT

We are very grateful to Kepa Ruiz-Mirazo for his collaboration in the past years.

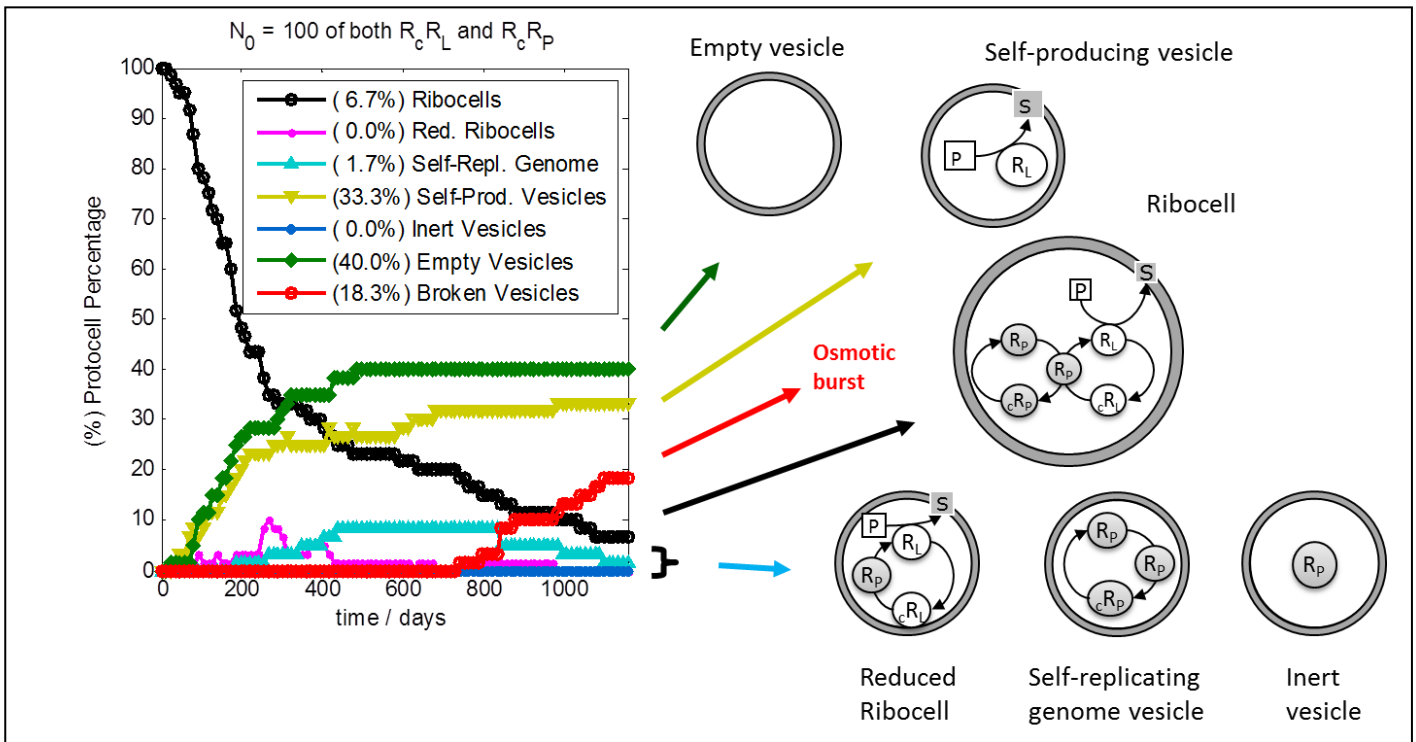

Fig. 10: Stochastic behavior of a population of 320 Ribocells: population composition against time (on the left), schematic draw of different protocells as result of vesicle division and random RNA strands distribution.

REFERENCES

[1] Walde, P. (ed.) 2005 *Prebiotic Chemistry from Simple Amphiphiles to Model Protocells. Topics in Current Chemistry*. Berlin:Springer.

[2] Chen, I. and Walde, P. (2010) From Self-Assembled Vesicles to Protocells. *Cold Spring Harb Perspect Biol* 2, a002170

[3] Luisi, P.L., Ferri, F. & Stano, P. 2006 Approaches to semi-synthetic minimal cells a review. *Naturwissenschaften* 93, 1-13.

[4] Luisi, P.L.and Stano, P. (2011) *The minimal cell. The biophysics of cell compartment and the origin of cell functionality*. Berlin: Springer.

[5] Noireaux, V. Maeda, Y.T. Libchaber A. (2011) Development of an artificial cell, from self-organization to computation and self-reproduction. *PNAS* 108, 3473

[6] Solé RV, Munteanu A, Rodriguez-Caso C and Macía J (2007) Synthetic protocell biology from reproduction to computation *Phil. Trans. R. Soc. B* 362, 1727-39.

[7] Stano, P.; Carrara, P.; Kuruma, Y.; Souza, T.; Luisi, P. L. (2011) Compartmentalized reactions as a case of soft-matter biotechnology: Synthesis of proteins and nucleic acids inside lipid vesicles. *J. Mat. Chem.* 21, 18887-902.

[8] Mavelli F. and Piotto S. (2006) Stochastic simulations of homogeneous chemically reacting systems. *J. Mol. Struct* 771, 55-64.

[9] Mavelli F. and Ruiz-Mirazo K. (2007) Stochastic simulations of minimal self-reproducing cellular systems, *Phil. Trans. R. Soc. B* 362, 1789-802.

[10] Stano, P., Wehrli, E. & Luisi P.L. 2006 Insights into the self-reproduction of oleate vesicles. *J. Phys: Condens. Matter* 18, 2231-2238.

[11] Mavelli F. and Ruiz-Mirazo K. (2013) Theoretical conditions for the stationary reproduction of model protocells. *Int. Biol.* 5 324-341.

[12] Mavelli F. and Ruiz-Mirazo K. (2011) ENVIRONMENT: a computational platform to stochastically simulate reacting and self-reproducing lipid compartments, *Phys. Biol.* 7, 036002.

[13] Ruiz-Mirazo K. and Mavelli F. (2008) On the way towards 'basic autonomous agents' stochastic simulations of minimal lipid-peptide cells. *BioSystems* 91, 374-87.

[14] Mavelli F. (2012) Stochastic simulations of minimal cells: the Ribocell model. *BMC Bioinformatics* 13(4), S10.

[15] Mavelli F. and Stano P. (2010) Kinetic models for autopoietic chemical systems role of fluctuations in homeostatic regime, *Phys Biol* 7, 016010.

[16] Gillespie DT (2007) Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem*, 58, 35-55.

[17] Varela F J, Maturana H R and Uribe R (1974) Autopoiesis: the organization of living systems, its characterization and a model. *Biosystems* 5, 187–96.

[18] Luisi P. L. and Varela F. J. (1990) Self-replicating micelles. A chemical version of minimal autopoietic systems, *Orig. Life Evol. Biosph.* 19, 633–43

[19] Zepik HH, Blöchliger E and Luisi P L 2001 A chemical model of homeostasis. *Angew. Chem., Int. Ed. Engl.* 40, 199–202.

[20] Takakura, K., Toyota, T. & Sugawara T. (2003) A Novel System of Self-Reproducing Giant Vesicles. *JACS* 125, 8134-8140

[21] Szostak J.W., Bartel D.P. and Luisi P.L., (2001) Synthesizing life. *Nature* 409, 387-90.

[22] Christensen U, (2007) Thermodynamic and kinetic characterization of duplex formation between 2'-O, 4'-C-Methylene-modified oligoribonucleotides, DNA and RNA. *Bioscience Reports*, 27:327-333

[23] Tsoi PY, Yang M, (2002) Kinetic study of various binding modes between human DNA polymerase and different DNA substrates by surface-plasmon-resonance biosensor. *Biochem. J.* 361: 317-325.

[24] Stage-Zimmermann T.K., Uhlenbeck O.C., (1998) Hammerhead ribozyme kinetics. *RNA*, 4: 875

[25] De Frenza A., (2009) Internal Technical Report, Chemistry Department University of Bari.

[26] Mansy S.S., Schrum J.P., Krishnamurthy M., Tobé S., Treco D.A., Szostak J.W., (2008) Template-directed synthesis of a genetic polymer in a model protocell. *Nature*, 454: 122-U10.

[27] Sacerdote M.G., Szostak J.W. (2005) Semipermeable lipid bilayers exhibit diastereoselectivity favoring ribose. *PNAS* 102, 6004-6008.

[28] Mavelli F., Della Gatta P., Cassidei G., Luisi P.L. (2010) Could the Ribocell be a Feasible Proto-cell Model? *OLEB* 40, 459-464.

[29] Mavelli F., (2011) Theoretical Approaches to Ribocell Modeling. In: *Minimal Cell*. Luisi P.L. and Stano P. (eds.), Dordrecht, Springer, 255-273.

[30] Smith T.F., Lee J.C., Gutell R.R., Hartman H. (2008) The origin and evolution of the ribosome. *Biol. Direct.* 3, 16-24.

[31] Agmon I. (2009) The Dimeric Proto-Ribosome: Structural Details and Possible Implications on the Origin of Life. *Int. J. Mol. Sci.* 10, 2921-2934.