From Never Born Proteins to Minimal Living Cells: Two Projects in Synthetic Biology

Pier Luigi Luisi · Cristiano Chiarabelli · Pasquale Stano

Abstract The Never Born Proteins (NBPs) and the Minimal Cell projects are two currently developed research lines belonging to the field of synthetic biology. The first deals with the investigation of structural and functional properties of *de novo* proteins with random sequences, selected and isolated using phage display methods. The minimal cell is the simplest cellular construct which displays living properties, such as self-maintenance, self-reproduction and evolvability. The semi-synthetic approach to minimal cells involves the use of extant genes and proteins in order to build a supramolecular construct based on lipid vesicles. Results and outlooks on these two research lines are shortly discussed, mainly focusing on their relevance to the origin of life studies.

Keywords cell-free protein expression · contingency · folding · liposomes · minimal cell · minimal genome · never born proteins · phage display · self-reproduction

Introduction

The term synthetic biology describes, rather broadly, those avenues of research, within life sciences, interested in the synthesis of parts of biological systems, or in the construction of models of biological systems. Synthetic biology comprises (and somehow is an extension of) biomimetic chemistry, but with the additional issue of "systems thinking."

Significant are the works of Joyce et al. dealing with the production of novel RNA species (Joyce and Orgel 1986; Santoro and Joyce 1997); or those by Benner, dealing with the synthesis of artificial peptides and nucleic acid analogs (Schneider and Benner 1990; Johnsson et al. 1993), or by Nielsen, on the peptide-nucleic acids (Nielsen 2004); as well as

---

P. L. Luisi (✉) · C. Chiarabelli · P. Stano
Biology Department, University of RomaTre, V.le G. Marconi 446, 00146 Rome, Italy
e-mail: luisi@mat.ethz.ch

C. Chiarabelli · P. Stano
Study Center "Enrico Fermi", Compendio del Viminale, 00184 Rome, Italy

the *de novo* production of proteins starting from a fix scaffold, as in the example given by Hecht (West et al. 1999). The works of von Kiedrowski (von Kiedrowski 1986) and Ghadiri (Lee et al. 1996) on self-replication of oligonucleotides and oligopeptides, respectively, represent an attempt to achieve synthetically the process of molecular replication, by designing a biologically inspired route. Finally, the Venter's proposal of inserting a synthetic genome in a bacterium (Zimmer 2003), or the study of semi-synthetic living cells in the laboratory, are all examples of research in this field. Many other examples could be given on the multifaceted and emerging issue of synthetic biology. The interested reader can refer to the recent review by Benner (Benner and Sismour 2005), or the editorial page by Hud and Lynn (Hud and Lynn 2004).

In this paper, we will present two projects from our group at the University of RomaTre, which can be rightly defined as belonging to the field of synthetic biology. The first one has to do with the synthesis of totally random, *de novo* proteins. As pointed out further on in this article, such proteins do not exist in nature, and they have been dubbed as "never born proteins" (NBPs). The second one has to do with the construction of cell models, in particular with minimal cells, defined as cells which contain the minimal and sufficient degree of complexity to be considered alive.

Following, we give a sketchy description of these two projects, referring to the original papers for more detailed analysis.

The Never Born Proteins (NBPs)

It is commonly accepted that the proteins existing in nature are only an infinitesimal part of the possible polypeptide sequences, and several ways of expressing this disparity have been formulated. One might say, for example, that the ratio between the possible and the actual protein sequences corresponds, in order of magnitude, to the ratio between the size of the universe and the size of a single hydrogen atom (for an discussion, see De Duve 2002). As trivial and old as it is, this consideration elicits some interesting questions about the origin of life, one of which being why and how these "few" extant proteins were selected during evolution.

A strict deterministic view would claim that the pathway leading to our proteins is more or less determined, and that there was practically no other viable alternative. According to this view, the historical route for the achievement of functional biopolymers, is somehow mostly determined, so as to permit the origin of life itself.

The contingency view would recite instead the following: if our proteins or nucleic acids have no special properties from a thermodynamic point of view, then *run the tape again* and a different "hydrogen atom" might have been produced – and perhaps the corresponding set of macromolecules would not necessarily have supported life.

At this point, some may say that proteins derive anyway from nucleic acid templates – perhaps through a primitive genetic code. But really this is no argument – as it merely shifts the problem of the aetiology of peptide chains to aetiology of nucleotide chains, while all the arithmetic problems remain more or less the same.

Looking at the vast realm of possibilities for non-existing proteins is challenging from a biochemical research point of view. What do all these NBPs look like? Are they only trivial variations of the proteins we already know, or – since their number is such an immensity – would it be possible that some of them possess unknown structure and properties? Have they not been produced simply and solely because of lack of time or bad luck – or due to the concomitance of some unknown and more subtle reason? (Of course we have many still

unknown proteins on Earth, but clearly the question of the non-selected NBPs has quite a different flavour).

A simpler question, related to structure, concerns folding: assuming we are able to make a large library of NBPs, what would be the frequency of folding – namely, which percentage of them would assume a stable tertiary conformation?

This question can be tackled by a concrete research project, which began recently in our group. The starting idea is to make a large number of NBPs, 50 amino acid residues long and test them in regards to tertiary folding. The particular question in mind is to measure the frequency of folding in the vast realm of NBPs.

By the well-known technique of phage display (Burton 1995; Hoess 2001), we produced a NBPs library that contains approximately 10^9 different polypeptide chains 50 residues long (the theoretical multiplicity of the sequence space for a peptide 50-mer is 20^{50}, i.e., 10^{65}, considering 20 possible amino acids). The phage-carried NBPs library was originated by the corresponding random DNA library (150 base-pairs long).

Figure 1a shows the construct obtained on the phage, with the *de novo* protein (NBP) fused with a protein of the phage capsid. In order to accomplish the process of selection, the primary random sequence of the NBP contains an internal -Pro-Arg-Gly- site that is cleavable by the proteolytic enzyme thrombin. In addition, for isolation purposes, the NBP was also engineered by adding a specific tag sequence (c-myc) at its N-terminus.

In other words, the genes of the library that was designed to perform this study were totally random except for the central nucleotides that codifies for the -Pro-Arg-Gly- internal sequence (Figure 1b).

The criterion utilized to probe the folding of the phage-displayed proteins is based on the resistance to the hydrolytic power of proteases (thrombin in the actual case), with the reasoning that folded chains are much more resistant than unfolded ones. This hypothesis, which is however quite reasonable, can be tested a posteriori, when the random proteins obtained by this strategy will be identified, produced, purified and their structure determined by spectroscopic methods.
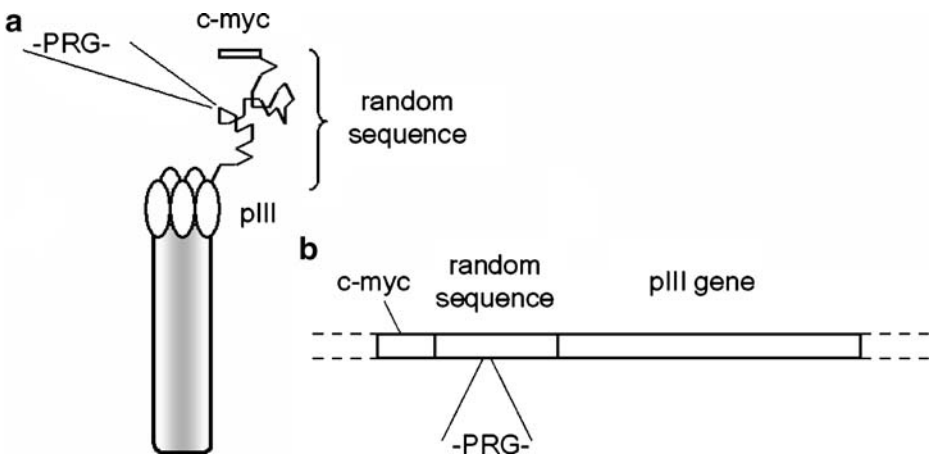
Figure 1 a The phage M13 with its five pIII capsid proteins. One of them is fused with the protein of interest (the NBP, a 50 amino acids long random sequence except for an internal -Pro-Arg-Gly- sequence), which is then displayed on the phage. In addition, at the N-terminus of such random sequence it is located the tag c-myc, which can be recognized by a specific antibody. b Detail of the plasmid used in the study. The plasmid carries three genes in sequence: c-myc gene, random gene (150 bp long), pIII gene.

Using the resistance to the action of thrombin as a criterion for the folding, and analysing a randomly taken sample of approximately 80 different phage constructs, it has been recently found (Chiarabelli et al., 2006a,b) that almost 5% of the NBPs displayed on the phage heads are rather resistant to the action of thrombin (see Figure 2).

If these preliminary data are confirmed by further analysis, a large fraction of the NBPs may fold into a stable tertiary structure. Therefore, folding is not a unique property of "our proteins," but a generalized one. As a corollary, since folding is a pre-requisite for functional biopolymers, the large number of folded chains present in a population of such random polypeptides (50-mers) could suggest that the probability of finding specific binding properties, or even catalysis, is relatively high.
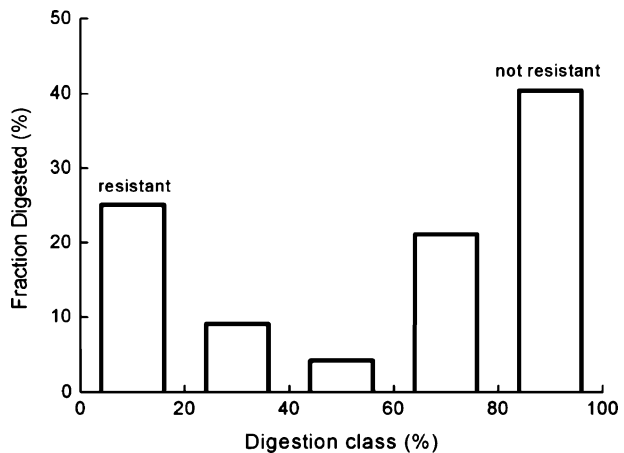
Seen in this light, contingency is associated with a relatively high probability for the (prebiotic) formation of folded proteins – which is positive in order to understand the origin of functional biopolymers. There is, however, another general teaching that we can draw out of these experiments. If "our" proteins are the product of contingency, then with all probability, the pathway to their prebiotic synthesis cannot be reproduced in the laboratory.

In conclusion, although it could be unfair, of course, to approach a problem of prebiotic chemistry by using sophisticated techniques of present-day molecular biology, in this particular case this strategy can actually provide an answer to the above question on the folding frequency of a vast library of random polypeptides. Refinement steps are possible and desirable, as well as the isolation and study of proteins, partly now in progress in our group.
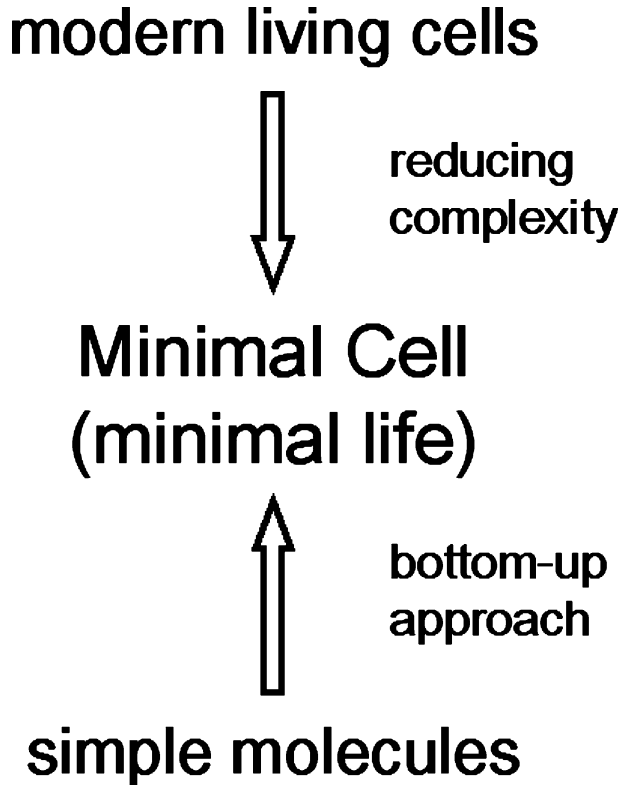
The Minimal Cell

It is assumed that cellular life on Earth originated from the inanimate matter, via an accretion of molecular and supramolecular complexity, up to the point where structures were produced, that had the novel and emergent property of being "living." With the aim of studying such processes, one can imagine to design experiments that largely reproduce this historical pathway (at the moment unknown), using step by step simulations of chemical and biochemical processes which occurred in the transition from non-living to living matter. This approach to simple – or "minimal" – cells, can be called "bottom-up," and suffers of the limitations described above: since the specific macromolecular sequences are the product of contingency, it will be impossible to design their synthetic pathway in the laboratory, and

Figure 2 Distribution of the peptide library with respect to thrombin digestion. About 25% of the phages used in the experiment carried a NBP that is not hydrolysed (or weakly hydrolysed) by thrombin.

Figure 3 The two approaches to the construction of the minimal cell.

modern living cells

↓ reducing complexity

Minimal Cell (minimal life)

↑ bottom-up approach

simple molecules

consequently, the pathway to living cells. It is legible to elicit the question on an alternative approach to the construction of minimal cells.

In Figure 3 we report a scheme that points out the two different approaches. The route going from extant cells to minimal ones uses extant nucleic acids and enzymes as components of a "reconstructed" minimal living cell, based on vesicle systems. In this way, the problem of the aetiology of specific macromolecular sequence is bypassed.

Whereas the term "bottom-up" is recognized and accepted, the terminology of this alternative route to the minimal cell is less clear. The term "top-down" was utilized in the past by our group (Luisi and Oberholzer 2001), mostly to set up a discrimination in comparison to the classic bottom-up approach. However, this terminology is not really correct, since in a way this is also a "bottom-up" approach, in the sense that it goes in the direction of increasing complexity (the living cell) starting from non-living components. Figure 4 shows this idea clearly; extant components are used, and synthetically assembled in a supramolecular construct: therefore this approach can be defined as "semi-synthetic."

In the following paragraphs, we outline the current research on minimal cells, making reference and reviewing the most recent achievements.

Reducing complexity

Looking at the simpler microbes, we realize that their cellular biology is based on thousands of expressed proteins that, more or less simultaneously, catalyse a myriad of reactions within the same tiny compartment – a maze of enormous complexity.

    Is such a complexity really essential for life – or might cellular life be possible with a much smaller number of components?

    Organisms such as *Mycoplasma genitalium* and *Buchnera* (an obligate parasite and an endosymbiont, respectively) are considered the simplest living cells, containing less than 500 coding regions. These organisms, however, live in highly permissive and biochemically rich conditions provided by their hosts.

    From the origin of life's point of view, we must consider that early cells could not have been as complex as our modern ones. Present enormous complexity is most likely the result of billions of years of evolution – with the development of a series of defence, repair and security mechanisms, as well as redundancies.
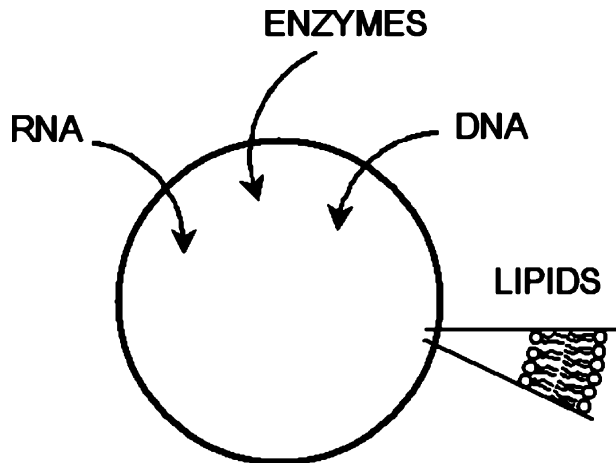
    Thus the general question "theoretically, how much can the structure of modern cells be simplified?" is related to the question about the structure of the early cells.

Defining minimal cells

All this brings to the notion of the minimal cell, defined as the one having the minimal and sufficient number of components to be called alive. What does "alive" mean? Living at the cellular level means the concomitance of three properties: self-maintenance (metabolism), self-reproduction, and evolvability.

    When all these three properties are fulfilled, we have full-fledged cellular life. Of course, in semi-synthetic systems the implementation is less than perfect, and several approximations to cellular life can be envisaged. For example we may have protocells capable of self-maintenance but not of self-reproduction; or vice versa. Or we can have protocells in which self-reproduction is active only for a few generations; or systems which do not have the capability of evolvability. And even in a given type of minimal cell – for example one with all three attributes – one might have quite different ways of implementation and sophistication.

Figure 4 The semi-synthetic approach to the construction of the minimal cell.

Minimal genomes

Organisms' complexity can be ranked by the size of their genome. The issue of "minimal genome" is related, of course, to the minimal cell. Going back to the above-mentioned *M. genitalium* and *Buchnera APS*, it is known that their genome size is of the order of 600 kb. In comparison, the most well studied bacterium, *E. coli*, has a genome of ca. 4,600 kb. Shimkets (Shimkets 1998) pointed out that the minimum genome size of a living cell should be around 600 kb; such claims have been discussed by Islas et al. (2004), which also provided a detailed analysis of a database of 641 proyariote genome sizes, discussing several aspects of minimal cell genome.

Taking into account the possibility of living in highly permissive conditions, Mushegian and Koonin (1996) calculated an inventory of 256 genes, which represents the amount of DNA required to sustain a modern type of minimal cell (for a critical commentary on these data, see Becerra et al. 1997). Recent studies further decreased this number to 206 genes (Gil et al. 2004), on the basis of comparative genomic analysis of endosymbionts and other microorganisms.

We can ask whether and at what extent it is possible to further reduce the number of genes by imagining a kind of theoretical knock-down of the genome that simultaneously reduces cellular complexity and part of non-essential functions (Luisi et al. 2002). This sort of analysis aims to focus on the essential functions, which confer the property of being alive.

A full discussion of this topic has been described in a recently published review (Luisi et al. 2006); here we will give only the most relevant points. For example, consider a semi-permeable cell, located in a highly permissive environment (rich in biochemically important compounds). Thanks to the external supply of low molecular-weight compounds, a minimal cell might live by just performing self-maintenance and self-reproduction. The key processes would be the self-reproduction of genes and proteins (the components of the "aqueous core" of the structure), and lipids (the "shell" of the structure), while modern important mechanisms, such as regulation, control and specific processing would not take place. The molecular machineries which carry out these tasks can be further simplified by introducing the idea of simplified protein synthesis and less-specific enzymes. Some indications suggest that ribosomal proteins may be not essential for protein synthesis (Weiner and Maizels 1987; Zhang and Cech 1998), and there are other suggestions about an ancient and simpler translation system (Nissen et al. 2000, Calderone and Liu 2004) and since a large part of the minimal genome is related to manipulation of nucleic acids, one can imagine that non-specific polymerases (maybe less efficient, but with a broader range of substrates) act as main nucleic acid processors (Luisi et al. 2002).

Finally, assuming a lower number of tRNA and amino acyl-tRNA synthase, we may arrive at the final estimation of ca. 50–100 genes as a simplified minimal genome (Luisi et al. 2006).

Of course, this sort of discussion takes us directly into the scenario of early cells at the origin of life, when the cells were probably "limping," but still potentially conserving the main properties of cellular life mentioned above.

A road map to the minimal cells

The analysis of minimal genome, shown in the previous section, still depicts a high level of complexity. At any rate, the necessary step to move towards the minimal cell is the entrapment of a minimal genome into a synthetic compartment.

Some important and concrete experimental advancements were reported in the last few years. According to Figure 4, the main subject of this research is the study of reaction within compartments, such as liposomes.

In this area, special attention is currently being devoted to the study of simple and complex biochemical reactions, e.g., protein expression, within lipid vesicles. The common strategy used by researchers operating in this field is to entrap all the components needed for protein expression within a compartment (generally lipid vesicles, but also water-in-oil emulsions), and follow the synthesis of a model (water soluble) protein, generally the green fluorescent protein (GFP).

In Table I we report a short description of the systems used by different authors. Oberholzer and Luisi (Oberholzer et al. 1999) described the first attempt to synthesize a polypeptide – poly(Phe) – within liposomes, and a couple of years later they reported on the synthesis of GFP (Oberholzer and Luisi 2002).

The Yomo and Urabe group, in Osaka, succeeded in the synthesis of GFP by incorporating cellular extracts into liposomes (Yu et al. 2001) and later on in the expression of a cascade network of two proteins, the T7 RNA polymerases and the GFP (Ishikawa et al. 2004).

Nakatani and Yoshikawa, on the other hand, were the first to report the synthesis of GFP in giant vesicles, i.e., micrometer-sized vesicles that can be observed using normal light microscope (Nomura et al. 2003).

A further step in this field was made in the work of Noireaux and Libchaber (Noireaux and Libchaber 2004). Their system deserves careful attention because – for the first time – the possibility of feeding the bioreactor by expressing a water soluble protein (α-hemolysin) that self-assembles into a lipid-soluble heptamer, which acts as a pore in the membrane, was taken into consideration. The particularly favourable cut-off of this pore (~3 kDa) allows small molecules enter into the vesicle core (thus feeding the reactor), keeping inside the high molecular-weight compounds that characterize the synthetic cell. The authors could demonstrate that such bioreactors could prolong the protein expression for ca. 100 h.

TABLE I Protein expression within compartments

Year	Authors	Main results
1999	Oberholzer et al.	Synthesis of poly(Phe) within lipid vesicles
2001	Yu et al.	Expression of GFP within lipid vesicles, as detected by flow cytometry
2002	Oberholzer and Luisi	Expression of GFP within lipid vesicles, followed by spectrofluorimetric analysis
2003	Nomura et al.	First report on protein synthesis within giant vesicles, higher internal yield of GFP in the first 3 h, when compared with the bulk synthesis, as evidenced by fluorescence microscopy
2004	Ishikawa et al.	Expression of two proteins: a genetic network in which the protein product of the first stage (T7 RNA polymerase) is required to drive the synthesis of the second protein (GFP), detected by flow cytometry
2004	Pietrini and Luisi	The expression of GFP is obtained by mixing four different water-in-oil compartments. The synthesis is followed by fluorescence microscopy
2004	Noireaux and Libchaber	α-Hemolysin, expressed within giant vesicles together with the reporter protein GFP, forms a membrane pore, that allows the feeding of the reactor in order to prolong the reactions for about 4 days

In Table I, we report also the use of water-in-oil macroemulsion as compartment for the synthesis of proteins. The advantage of this system, in comparison to liposomes, is related to the high entrapment yield and the possibility of fusion or solubilisate exchange between different compartments. In this way, it was possible to observe protein expression after mixing different compartments, each containing a sub-set of the components needed to perform the reaction (Pietrini and Luisi 2004).

Although limited in number, the handful of paper on protein expression within compartments encourage to proceed further and face the problem of replication of all the "core" components (DNA, RNA, proteins), or the shell reproduction by internal synthesis of the membrane-forming compound. At this stage, we must realize that – in contrary to the studies reported in Table I – it is important to know the exact number of macromolecular components needed for protein expression. In fact, replicating the core components means that the synthesis of all the enzymes for protein expression, the genetic material replication (which can be DNA or even mRNA), the ribosomal protein and rRNA replication, and the tRNA replication must all be achieved. The works done up till now employed commercially available cellular extracts (mainly from *E. coli*), which are a sort of black-box. Very recently, the group of Ueda, in Tokyo, introduced a cell-free transcription and translation system composed by purified components (Shimizu et al. 2001). This kit, now commercially available with the trade name of PURESYSTEM®, will certainly be functional to the studies on minimal cell. It perfectly fits the requirement of a full synthetic biology approach, providing a route for the assembling of minimal cells with a known and adjustable chemical composition.

Going back to the original question, the replication of all the components inside a minimal cell requires, in addition to the expression of all the components constituting the protein synthesis machinery, the expression of the (several) proteins that replicate the DNA itself. Considering that – theoretically – an RNA genome instead of the corresponding DNA genome can be implemented, an alternative route could involve the use of RNA polymerase enzymes (for example Qβ-replicase, that has been already used for compartmentalized reactions, as well as other simpler RNA-dependent RNA polymerases) as a minimal enzyme set that accomplishes the genetic material replication.

The biosynthesis of lipids from within a lipid vesicle is another ambitious task for the advance in the route map to the minimal cell. The idea is rather simple: a synthetic compartment (delimited in space by its boundary) should be able to produce – from within, and in addition to the replication of the inner components (DNA, proteins, etc.) – the components that form the boundary itself. In other words a core-and-shell coupled reproduction is necessary.

In 1991, Luisi et al. (Schmidli et al. 1991) reported the first attempt to use a lipid-synthesizing lipid vesicle. The approach involved the incorporation of the enzymes that produce lipids (i.e., lecithin) in lipid vesicles. The four enzymes of the lecithin salvage pathway were reconstituted in lecithin vesicles, then water-soluble precursors were added, and enzymatically transformed *in situ* into newly synthesized lecithin molecules. In terms of modern molecular biology approach to minimal cells, the expression of membrane enzymes (as those that synthesize lipids) in lipid vesicles represents an important goal; a recent work reports the cell-free synthesis of membrane protein (Kuruma et al. 2005).

As soon as this technical problem is solved, one might imagine a liposome that produces, from the within, the components of its lipid membrane. Due to the increase of surface, the liposome would grow in size and possibly split in two or more liposome-daughters (Figure 5). In this highly simplified scheme, a synthetic cell containing all the genes, enzymes, tRNAs and ribosomes, performs the reproduction of all the components

Figure 5 A cell that makes its own boundary. The complete set of biomacromolecules needed to perform protein synthesis (genes, RNA polymerises, ribosomes, tRNA and other enzymes) is indicated as *Rib*. The product of this synthesis (indicated as *E*) is the complete set of enzymes for lipid (*L*) synthesis, which use the precursor A. The cell can use all the low molecular-weight precursors and energy rich compounds (*B*, supposed available, for example deoxynucleotides, nucleotides, amino acids, lipid precursors and all the low molecular-weight compounds) in order to accomplish the reproduction of the whole machinery Rib and to sustain biochemical reactions. In order to avoid the "death by dilution," i.e., the generation of new cells that miss at least one cellular component, the cell should perform a core- and shell-reproduction; moreover, the components should distribute evenly between the new compartments in the act of division.
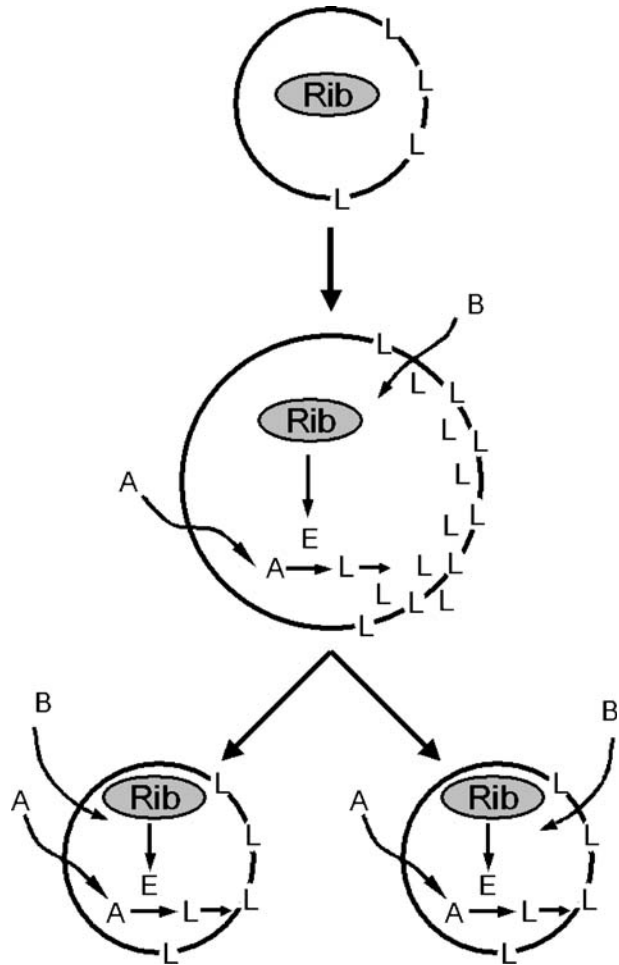
(DNA, RNA, proteins) of the aqueous core as well as the lipids, thanks to the external supply of deoxynucleotides, nucleotides, amino acids, lipid precursors and all the low molecular-weight compounds needed to accomplish these biosynthesis. This model represents a molecular construct that self-replicates from the within, i.e., an autopoietic semi-synthetic cell.

Since the processes is not regulated, in order to avoid the so-called "death by dilution," i.e., the absence of at least one of the key macromolecular compounds in the progeny of a minimal cell, the core- and shell-reproduction should occur with similar rates, and the solutes should distribute evenly between the newly formed cells, as shown in the ideal case of Figure 5.

Concluding Remarks

The NBPs and minimal cell projects are currently under progress. The main goal of the first project is selection and production of the first *de novo* proteins, whose structures must be

determined and whose biological activity explored (Chiarabelli et al., 2006a,b). On the other hand, the minimal cell project follows the perspectives outlined previously, improving the conditions for protein expression and, in particular, looking for the conditions for expression of lipid-soluble proteins. Furthermore, the reduction of the number of components needed for protein expression is also important regarding the issue of minimal genome and self-reproduction of the components constituting a minimal cell.

References

Becerra A, Islas S, Leguina JI, Silva E, Lazcano A (1997) Polyphyletic gene losses can bias backtrack characterizations of the cenancestor. J Mol Evol 45:115–118

Benner SA, Sismour AM (2005) Synthetic biology. Nat Genet 6:533–543

Burton DR (1995) Phage display. Immunotechnology 1:87–94

Calderone CT, Liu DR (2004) Nucleic-acid-templated synthesis as a model system for ancient translation. Curr Opin Chem Biol 8:645–653

Chiarabelli C, Vrijbloed JW, Thomas RM, Luisi PL (2006a) Investigation of *de novo* totally random biosequences. Part I. A general method for *in vitro* selection of folded domains from a random polypeptide library displayed on phage. Chem Biodivers 3:827–839

Chiarabelli C, Vrijbloed JW, De Lucrezia D, Thomas RM, Stano P, Polticelli F, Ottone T, Papa E, Luisi PL (2006b) Investigation of *de novo* totally random biosequences, Part II. On the folding frequency in a totally random library of *de novo* proteins obtained by phage display. Chem Biodivers 3:840–859

De Duve C (2002) Life evolving: molecules, mind, and meaning. Oxford University Press, Oxford, UK

Gil R, Silva FJ, Peretó J, Moya A (2004) Determination of the core of a minimal bacteria gene set. Microbiol Mol Biol Rev 68:518–537

Hoess RH (2001) Protein design and phage display. Chem Rev 101:3205–3218

Hud NV, Lynn DG (2004) Model systems: from life's origins to a synthetic biology. Curr Opin Chem Biol 8:627–628

Ishikawa K, Sato K, Shima Y, Urabe I, Yomo T (2004) Expression of a cascading genetic network within liposomes. FEBS Lett 576:387–390

Islas S, Becerra A, Luisi PL, Lazcano A (2004) Comparative genomics and the gene complement of a minimal cell. Orig Life Evol Biosph 34:243–256

Johnsson K, Allemann RK, Widmer H, Benner SA (1993) Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides. Nature 365:530–532

Joyce GF, Orgel LE (1986) Nonenzymatic template-directed synthesis on RNA random copolymers – poly (C,G) templates. J Mol Biol 188:433–441

Kuruma Y, Nishiyama K, Shimizu Y, Muller M, Ueda T (2005) Development of a minimal cell-free translation system for the synthesis of presecretory and integral membrane proteins. Biotechnol Prog 21:1243–1251

Lee DH, Granja JR, Martinez JA, Severin K, Ghadiri MR (1996) A self-replicating peptide. Nature 382:525–528

Luisi PL, Oberholzer T (2001) Origin of life on Earth: molecular biology in liposomes as an approach to the minimal cell. In: Giovanelli F (ed) The bridge between the big bang and biology. CNR, Rome, Italy, pp 345–355

Luisi PL, Oberholzer T, Lazcano A (2002) The notion of a DNA minimal cell: a general discourse and some guidelines for an experimental approach. Helv Chim Acta 85:1759–1777

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 93:1–13

Mushegian A, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci USA 93:10268–10273

Nielsen PE (2004) PNA technology. Mol Biotechnol 26:233–248

Nissen P, Hansen J, Ban N, Moore PB, Steitz TA (2000) The structural basis of ribosome activity in peptide bond synthesis. Science 289:920–930

Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. Proc Natl Acad Sci USA 101:17669–17674

Nomura SM., Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. ChemBioChem 4:1172–1175

Oberholzer T, Nierhaus KH., Luisi PL (1999) Protein expression in liposomes. Biochem Biophys Res Commun 261:238–241

Oberholzer T, Luisi PL (2002) The use of liposomes for constructing cell models. J Biol Phys 28:733–744

Pietrini A, Luisi PL (2004) Cell-free protein synthesis through solubilisate exchange in water/oil emulsion compartments. ChemBioChem 5:1055–1062

Santoro SW, Joyce GF (1997) A general purpose RNA – cleaving DNA enzyme. Proc Natl Acad Sci USA 94:4262–4266

Schmidli PK, Schurtenberger P, Luisi PL (1991) Liposome mediated enzymatic synthesis of phosphatidyl-choline as an approach to self-replicating liposomes. J Am Chem Soc 113:8127–8130

Schneider KC, Benner SA (1990) Oligonucleotides containing flexible nucleoside analogs. J Am Chem Soc 112:453–455

Shimkets LJ (1998) Structure and sizes of genomes of the archaea and bacteria. In: De Bruijn FJ, Lupskin JR, Weinstock GM (eds) Bacterial genomes: physical structure and analysis. Kluwer, Boston, MA, pp 5–11

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755

Von Kiedrowski G (1986) A self-replicating hexadeoxynucleotide. Angew Chem, Int Ed Engl 25:932–935

Weiner AM, Maizels N (1987) tRNA – like structures tag the 3′ ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. Proc Natl Acad Sci USA 84:7383–7387

West MW, Wang WX, Patterson J, Mancias JD, Beasley JR, Hecht MH (1999) De novo amyloid proteins from designed combinatorial libraries. Proc Natl Acad Sci USA 96:11211–11216

Yu W, Sato K, Wakabayashi M, Nakatshi T, Ko-Mitamura EP, Shima Y, Urabe I, Yomo T (2001) Synthesis of functional protein in liposome. J Biosci Bioeng 92:590–593

Zhang B, Cech TR (1998) Peptidyl-transferase ribozymes: trans reactions, structural characterization and ribosomal RNA-like features. Chem Biol 5:539–553

Zimmer C (2003) Tinker, tailor: can venter stitch together a genome from scratch? Science 299:1006–1007