

**REVIEW****Chemical Aspects of Synthetic Biology**by **Pier Luigi Luisi**Department Biology, University of Roma TRE, Viale G. Marconi 446, I-00146 Roma  
(phone/fax: +390655176329; e-mail: luisi@mat.ethz.ch)

---

Synthetic biology as a broad and novel field has also a chemical branch: whereas synthetic biology generally has to do with bioengineering of new forms of life (generally bacteria) which do not exist in nature, 'chemical synthetic biology' is concerned with the synthesis of chemical structures such as proteins, nucleic acids, vesicular forms, and other which do not exist in nature.

Three examples of this 'chemical synthetic biology' approach are given in this article. The first example deals with the synthesis of proteins that do not exist in nature, and dubbed as '*the never born proteins*' (NBPs). This research is related to the question why and how the protein structures existing in our world have been selected out, with the underlying question whether they have something very particular from the structural or thermodynamic point of view (for example, the folding). The NBPs are produced in the laboratory by the modern molecular biology technique, the phage display, so as to produce a very large library of proteins having no homology with known proteins.

The second example of chemical synthetic biology has also to do with the laboratory synthesis of proteins, but, this time, adopting a prebiotic synthetic procedure, the fragment condensation of short peptides, where short means that they have a length that can be obtained by prebiotic methods; for example, from the condensation of *N*-carboxy anhydrides. The scheme is illustrated and discussed, being based on the fragment condensation catalyzed by peptides endowed with proteolytic activity. Selection during chain growth is determined by solubility under the contingent environmental conditions, *i.e.*, the peptides which result insoluble are eliminated from further growth. The scheme is tested preliminarily with a synthetic chemical fragment-condensation method and brings to the synthesis of a 44-residues-long protein, which has no homology with known proteins, and which has a stable tertiary folding.

Finally, the third example, dubbed as '*the minimal cell project*'. Here, the aim is to synthesize a cell model having the minimal and sufficient number of components to be defined as living. For this purpose, liposomes are used as shell membranes, and attempts are made to introduce in the interior a minimal genome. Several groups all around the world are active in this field, and significant results have been obtained, which are reviewed in this article. For example, protein expression has been obtained inside liposomes, generally with the green fluorescent protein, GFP. Our last attempts are with a minimal genome consisting of 37 enzymes, a set which is able to express proteins using the ribosomal machinery. These minimal cells are not yet capable of self-reproduction, and this and other shortcomings within the project are critically reviewed.

---

**Introduction.** – I am a chemist who left his original avenues of polymer chemistry to move towards biochemistry and biology. Reflecting on the research field I am now pursuing, I can, however, see that somehow the 'genes' of chemistry have remained, and they actually have helped me to move at the interface between biology and chemistry. Thus, the projects that I will describe below can be categorized within the

broad and novel field of synthetic biology, but they are all have a strong character of chemistry.

In fact, this novel and fashionable term ‘*synthetic biology*’ is now used to indicate a field where generally existing life forms are modified and the genomic content re-directed towards novel, non-existing life forms; for example, bacterial life that does not exist on earth. Examples of these techniques can be found in recent issues in *Nature* [1] and *Science* [2]. This is all based on the hard hand of the bio-engineering approach that thrives from classic DNA molecular biology. The chemical approach to synthetic biology I am talking about is one that, instead of hampering with living life forms and creating some more or less fortunate imitations, aims more simply at the synthesis of *molecular structures* and/or *multi-molecular organized systems* that do not exist in nature. These man-made, in nature non-existing biological molecular or supramolecular structures can be obtained either by chemical or biochemical synthesis, possibly with the help of mechanical manipulations thereof.

I would put in this category of *chemical synthetic biology* the well-known work of *Albert Eschenmoser* on nucleic acids containing pyranose instead of ribose [3], structures that have been synthesized in the laboratory, and that do not exist in Nature. The question is possibly why Nature did not make them, and much can be learned from the very asking of this question. The chemical modifications of nucleic acid bases pursued by *Steve Benner* [4] belong also to this class of studies.

There are examples also in the field of proteins: the synthesis of proteins containing a reduced alphabet – only 3, 5, 7, or 9 amino acids – already described in the literature [5][6] belong, in my opinion, to this field of ‘*chemical synthetic biology*’. Also the approach pioneered by *Craig Venter* and co-workers, aimed at synthesizing an entire genome by chemical methods [7], can be considered as one of the examples of chemical synthetic biology.

In the following, I would like to present three projects carried out in my laboratory that can be considered as also belonging to this chemical frontier of synthetic biology. One project is carried out under the name of ‘*the never born proteins*’, meaning proteins that have not been produced and/or selected by nature in the course of biological evolution. This synthetic procedure also produces the corresponding ‘never born m-RNA’.

The second project deals with the synthesis of specific macromolecular sequences by fragment condensation under simulated environmental pressure which corresponds to molecular evolution.

The third is the ‘*minimal cells*’ project, meaning semi-synthetic cells that do not exist in nature, which may represent the simplest form of cellular life.

These three projects are at different degrees of progress in my laboratory, and I will describe their present stage and outlook.

**1. The ‘Never Born Proteins’ (NBPs).** – The starting point is the numerology of proteins, in particular the well-known consideration that the proteins existing in nature make only an infinitesimal fraction of the theoretically possible structures. This paradox has been emphasized by various authors, also by *Christian de Duve* in his latest book [8]. There are many ways to express this. For example, one can say that the ratio between the theoretically possible proteins having a chain with 100 residues, and the actual number of all existing proteins (probably something around  $10^{14}$ ), comes close to

the ratio between the space of the universe and the space of a single H-atom; or, using a more earthly example, close to the ratio between the all sand of Sahara and one single grain of sand [9].

These astronomic figures may appear deprived of practical physical meaning. However, they convey a very simple, well graspable concept, *i.e.*, that our life is based on a very limited number of structures; and this, in turn, elicits a very relevant question: how and why have these few structures been selected out?

1.1. *The Different Viewpoints: Determinism vs. Contingency.* There are different answers one can tentatively give to this last question.

One first possible answer is that ‘*our*’ proteins have something very special that made the selection possible. For example, they might be the only ones to be stable; or water soluble; or those which have very particular viscosity and/or rheological properties. In all these cases, they would have been selected because of their particular physical properties.

A second point of view is that our proteins have no extraordinary physical properties at all; they have been selected by chance among an enormous number of possibilities of quite similar compounds. They came out by ‘*chance*’, and it happened that they were capable of fostering cellular life. The term ‘*chance*’ is nowadays substituted by the more elegant term ‘*contingency*’. Cast the dice again, and the probability that exactly our  $10^{14}$  or so proteins come out again is at all effects practically zero, so that life as it is now may not have started. One can conceive some different forms of life thriving on quite different proteins, but this remains to be established.

Of course, contingency never works alone, it is always accompanied by some deterministic laws – certainly by thermodynamics and energy minimization processes – but, according to the contingency view, basically ‘*our*’ life would be a serendipitous property of these casually determined structures.

The deterministic view mentioned above may, instead, assume some alternative extreme positions, up to the point of saying that life is an inescapable outcome of the laws of nature, and that, therefore, all prerequisites for making life, including the basic macromolecular structures, are determined.

In this sense, an author who should be particularly kept in mind is *Christian de Duve*, who in his book [8] stated:

*‘It is self-evident that the universe was pregnant with life and the biosphere with man. Otherwise, we would not be here. Or else, our presence can be explained only by a miracle...’*

This is basically the view that the origin of life was an obligatory, inevitable process, and if one literally takes this view, then one has to conclude that the proteins must have been chosen in the right way so as to make life possible. One cannot, in fact, assume the inevitability of life and then let contingency shape the structure of proteins as chance structures.

To me, as I expressed in my recent book [9], the view that life is inescapable corresponds to implying a form of intelligent designer, and this, as the Anglican priest *Paley* said hundreds of years ago [10], cannot be else than God. In fact, I dubbed the authors that adhere to this view, including those of the anthropic principle [11][12] ‘*crypto-creationist*’ [9] (not to be confused, however, with the American creationism, which is simply a form of fundamentalism). The view of *contingency* in evolution and

life is advocated, among many others, most notably by *Stephen J. Gould* [13] and *J. Monod* [14].

1.2. *The Experimental Project.* The basic idea of the project is to test whether ‘our’ proteins have really something particular with respect to the proteins that have never existed. How can one conduct this project? Simply by synthesizing proteins that do not exist in nature, and comparing them with ‘our’ proteins. It is a project of chemical synthetic biology, as outlined in the *Introduction*, aimed at producing a quite different ‘*grain of sand*’, which should be the product of random choice – and asking then the question, whether ‘our’ proteins are really so different and peculiar with respect to those synthetic biology products – in terms of stability, solubility, or folding. Actually, folding is a particularly important and stringent criterion, as the prerequisite for the biological activity of proteins is their globular folding, which is a consequence of the primary structure.

Such a project has been initiated at the Federal Institute of Technology in Zürich, Switzerland, to be pursued by my group transferred to the University of Rome3, Italy, in particular, by *Cristiano Chiarabelli* and *Davide de Lucrezia*. The first set of papers describing these results about the ‘*never born proteins*’ (NBPs) has been recently published [15a–d].

The principle to produce NBPs is simple: if one makes a long string (say 150 bases) of DNA purely randomly, the probability of hitting an existing sequence in our Earth is practically zero (it corresponds to a number equivalent to the ratio between one grain of sand and the entire Sahara). If you then let this DNA being processed by standard recombinant DNA and *in vivo* expression techniques, you will obtain a 50-residues-long polypeptide that does not exist on Earth, and when this polypeptide is globularly folded, you have already obtained a NBP.

In practice, what we do is to work with a large library of DNA by the so-called phage-display method. We obtain first from commercial sources a library of totally random DNA sequences with the desired length (150 base pairs in our case). The random DNA segment is then inserted within a phage genome so that the corresponding random protein is linked to a capsid protein. The production of the phage library actually needs the infection of cells that provide the machinery for the synthesis of viral proteins. Those proteins will be displayed on the capsid of the phage (one per phage), and they are (in the N-termini portion) totally random, *de novo* proteins. In our case, the sequence of the NBPs is not completely random, since a tripeptide sequence has been inserted in the middle of the random sequence with the aim of selection (*vide infra*).

This is the basis of the work carried out by us [15]. In this way, by a first run, a library of *ca.*  $10^9$  of 50-residues-long polypeptides was obtained. The first questions at this point were: *i*) are they really all ‘*never born proteins*’, *i.e.*, more specifically, are they really absent in the protein data bank collected until now? and: *ii*) what will be the fraction of folded polypeptides, *i.e.*, of globular proteins?

It is also clear that, for practical reasons, one cannot study all  $10^9$  clones; one can only refer to a selected sampling of it, chosen, however, without any preconceived bias so that the statistical relevance of their properties still hold.

Let us begin with the question about being ‘*never born*’. The 79 sequences which were selected at random were compared with known protein sequences, and no

similarity was found, although a permissive criterion was adopted for the comparative analysis [15].

In conclusion, then, the proteins so synthesized can indeed be considered as non-extant, which permits the terminology of ‘*never born proteins*’ (NBPs). Of course, it is possible that some of these sequences may have been proposed in the course of molecular evolution, and then gone lost; or that some of them are present in some unexplored plants or micro-organisms of our Earth. But, in first and good approximation, they are not present in any living form we know.

The other question (about folding) has been tackled based on the well-accepted observation that folded proteins are not easily digestible by proteases. The strategy involved the insertion of the tripeptide PRG (proline-arginine-glycine), substrate for the proteolytic enzyme thrombin, in the otherwise totally random protein sequence (*i.e.*, the DNA library was designed in order to have three non-random codons in the middle of the sequence). In this way, each of the new proteins had the potentiality of being digested by the enzyme, with the expectation, however, that globularly folded NBPs would be protected from digestion. With this idea in mind, the 79 randomly selected clones were incubated in a medium in the presence of thrombin. The larger part of the population was rapidly hydrolyzed, but *ca.* 20% of the population was highly resistant to the action of thrombin.

Although our criterion of folding should be considered at this point is an approximate one, 20% is a surprisingly high figure. It suggests that folding is indeed a general property, something that arises naturally, even for proteins of medium length.

The characterization of some of those folded proteins has begun, and, in *Fig. 1*, the circular dichroic properties in the far UV region of two of them, labeled preliminarily as **A** and **B**, are shown (for the primary sequence, see the original reference [15]). It is apparent that, in both, a significant percentage of periodic structure,  $\alpha$ -helix in particular, is present, and, furthermore, very interestingly, the globular folding is thermoreversible, indicating that is under thermodynamic control.

In *Fig. 2*, the computed folding [15b] of these two proteins are illustrated, according to the analysis carried out by Dr. *Fabio Politicelli* in Roma<sup>3</sup>. Although the computational method used (Rosetta) is the most reliable of present day’s literature, such three-dimensional drawings cannot be considered yet as the definitive structure; for the actual structure, one should await for NMR or X-ray data.

We have now about one dozen of such computed structures, and, although one should wait before attempting generalizations, it appears rather safe at this point to state that folding and thermodynamic stability are not properties that are restricted to our extant proteins, and that, on the contrary, they appear to be rather common features of randomly created polypeptides.

On the basis of this, one is tempted to propose that ‘our’ proteins do not belong to a class of polypeptides with privileged physical properties. And, by inference, one could say that this kind of data, once confirmed by a larger number of cases, permit to brake a lance in favor of the scenario of contingency.

Of course, the NBPs may also have bio-technological importance, and may be also very interesting from the structural point of view: could they, for example, display novel catalytic and structural features that have not been observed in ‘our’ proteins? The answer to these question must await for much more data.

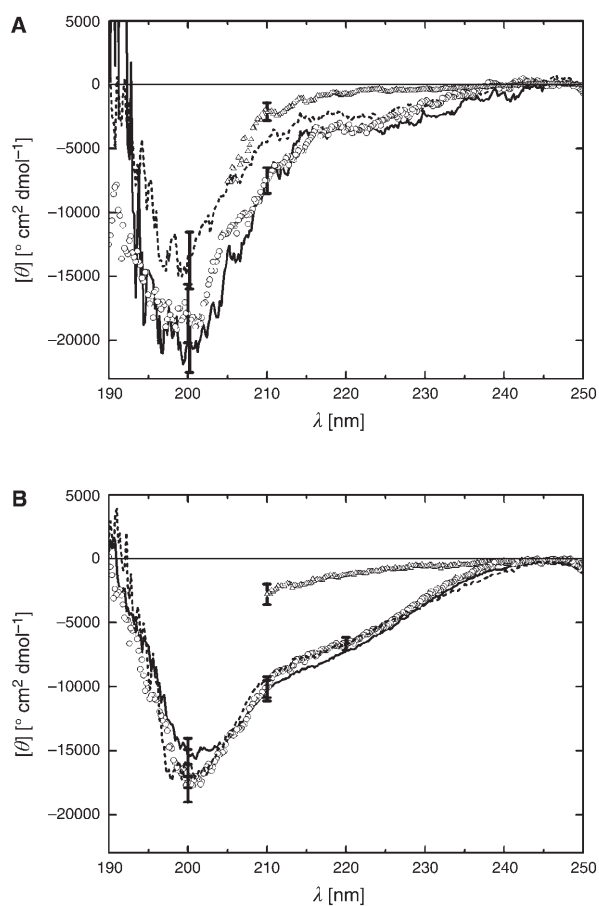


Fig. 1. CD Spectrum of two 'never born proteins'. Note the significant content of secondary structure and the reversibility of the folding with temperature. For a detailed description of these CD experiments, see [15b]. The predicted tertiary structure of the two proteins is shown in Fig. 2.

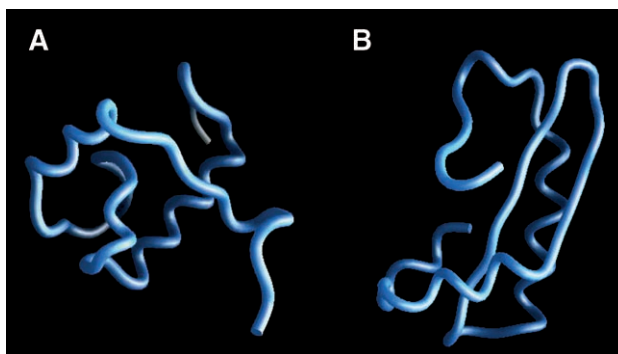


Fig. 2. The structure prediction of the two 'never born proteins' firstly characterized. The predicted structures appear to be qualitatively in agreement with the spectroscopic data of CD and fluorescence.

1.3. *The Case of Never Born RNAs.* There is an important addition to be made to the above synthetic-biology program: this is that the synthesis of NBPs is automatically accompanied by the synthesis of the corresponding m-RNAs. This permits to tackle the question whether and to what extent such totally random RNAs are going to be folded. We have conducted an analysis of several randomly chosen ‘never born’ RNAs [15c][15d], and found indeed an extensive folding, which we could partly classify in different classes by utilizing an *ad hoc* developed method of analysis, based on a nuclease enzyme ( $S_1$ ) coupled with a temperature gradient (the ‘Foster assay’; see [15c]). Particularly interesting was the observation of an RNA structure which did not unfold at temperatures as high as 60°. This led us to the hypothesis that thermo-resistant RNA structures may not be so rare.

Now, we are developing the work on NBPs and corresponding RNAs in two directions: one is the characterization of the already made NBPs: we would like to obtain a large number of NBP structures so as to have a statistically significant display. The other direction of work is to prepare another library of NBPs, this time with a length of 20 residues. In this case, we will be mostly looking for primitive forms of catalysis, which, given the small size, is not expected to be exceptional, but relevant for the origin of life (see the following project). A length of 20 amino acid residues corresponds to a corresponding RNA length of 60 nucleotides, and this is a particularly interesting size, *i.e.*, close to most of the ribozymes’ sizes.

**2. Synthesis of Polypeptides under Simulated Prebiotic Evolution.** – 2.1. *The Status of the Matter.* The previous Section addresses the question of the frequency of foldable chains in a library of totally random *de novo* polypeptides, whereby such chains have been obtained by modern molecular-biology techniques. In this respect, we were interested in the properties of such polypeptides, and not on the chemistry of their formation under prebiotic conditions.

Therefore, one of the main questions about the origin of macromolecules remains open: how have *multiple copies* of *identical* long specific chains been produced? Again, that the polypeptides came from long nucleic acids is not an answer, as the question would then be referred to the etiology of specific sequences of polynucleotides.

The synthesis of long homo-polypeptides or homo-polynucleotides, *i.e.*, chains containing only one type of residue [16a–c][17], has been described, but this does not solve the problem. The problem is the synthesis of co-oligopeptides, *i.e.*, chains containing different amino acid residues (or nucleotides), and it is well-known from standard theory of copolymerization that the synthetic procedures valid for homopolymers are generally not applicable to the synthesis of a mixture of co-monomers, and that the monomer composition in the copolymer can be significantly different from that in the starting monomer mixture [9]. Furthermore, even if all amino acids present in the mixture would be polymerized with the same probability – the case of an ideal copolymerisation – one would obtain copolymers with a random distribution of residues along the chain, which is not what we want. A method that produces long copolymers with random composition is the one used by Fox and co-workers [18], which also – aside from the problem of the lack of characterization of these compounds – does not solve the problem of the synthesis of identical chains.

In fact, if one searches in the literature for the prebiotic syntheses (*Merrifield* method excluded) of relatively long co-oligopeptides (say at least 30 residues, so that they partly begin to assume a stable folding), one finds almost nothing. Some references are collected in [9].

The group of *Auguste Commeyras* has approached the problem of the prebiotic formation of peptides by using the condensation of *N*-carboxy anhydrides (NCA) [19][20], a method that, according to the authors, is prebiotic; but also in this case the critical question of the production of multiple identical copies of long (30 residues or more) co-oligopeptides could not yet be achieved.

How then can one conceive a co-polymerization scheme which produces, for example, lysozyme-kind of molecules? This question forms the basis of our next chemical synthetic-biology project.

*2.2. The Underlying Model.* We need first a work hypothesis for the formation of multiple copies of identical long co-oligopeptide chains. One such hypothesis is contained in our research project, conducted in collaboration with *Peter Strazewski* and *Peter Goekjian* at the University of Lyon, France. The basic idea is that such a chain elongation proceeds by successive *fragment condensation* of prebiotically formed short co-oligopeptides (*i.e.*, peptide-bond formation, *i.e.*, the reverse reaction of the peptide bond hydrolysis) as indicated in [9]. In particular, the synthesis of short peptides is realized by the prebiotic NCA condensation. A key assumption is that, in this random library, some peptides may arise, which possess proteolytic activity. Further, one assumes that fragment condensation may be induced by the catalytic action of such peptides.

How realistic are these two assumptions? It has been already reported that even simple peptides may be endowed with proteolytic activity. For example, His–Ser appears to be capable of cleaving peptide and nucleic acid bonds [21]; and even Gly–Gly [22] appears to possess some catalytic activity.

Thus, the idea that a random family of peptides containing Ser and His may possess proteolytic activity is not so unreasonable. In our case, we need, however, the reverse reaction, *i.e.*, the synthesis of peptide bonds. Here again, it is known that, in principle, proteolytic enzymes are capable of inducing peptide-bond formation. Extensive review articles have been presented in the past by *Jakubke et al.* [23], and by others, including our own group [24]; and, within the field of the origin of life, scenarios of alternate dry and wet environments have been theoretically proposed as conditions for bond-formation and chain elongation [20].

One may then consider to start from a prebiotic library of, say, decapeptides (this length is quite possible with the NCA method) and proceed with fragment condensation induced by catalytically active peptides.

It should be taken into consideration that the random condensation of all partners of a medium-size library of co-oligopeptides, after a few condensation steps, would give rise to an astronomic number of longer chains. The selection of only one or very few chain configurations out of this random library is possible only in the presence of some stringent selection criteria. Which selection criteria might have been possible in the prebiotic scenario? Clearly, only those based upon chemo-physical properties and chemo-physical conditions.

Thus, we arrive at another key assumption of our working program: The idea is that the selection is governed by the contingency of the environmental conditions, such as



pH, solubility, temperature, salinity, *etc.* Contingently upon these conditions, the largest majority of the library structures may be eliminated (*e.g.*, by lack of solubility, or due to aggregation), and only a few chain products may ‘*survive*’ in solution, undergoing then further elongation in solution. Thus, the selection criteria conceived in our work is one that is assumed to simulate the natural chemical evolution – in particular a kind of survival of the best fit as governed by the interplay of contingent conditions (the actual pH or salinity or temperature operating at that moment of growth) and the actual physical properties of the candidate chains. Fig. 3, taken from [9], gives an illustration of this process.

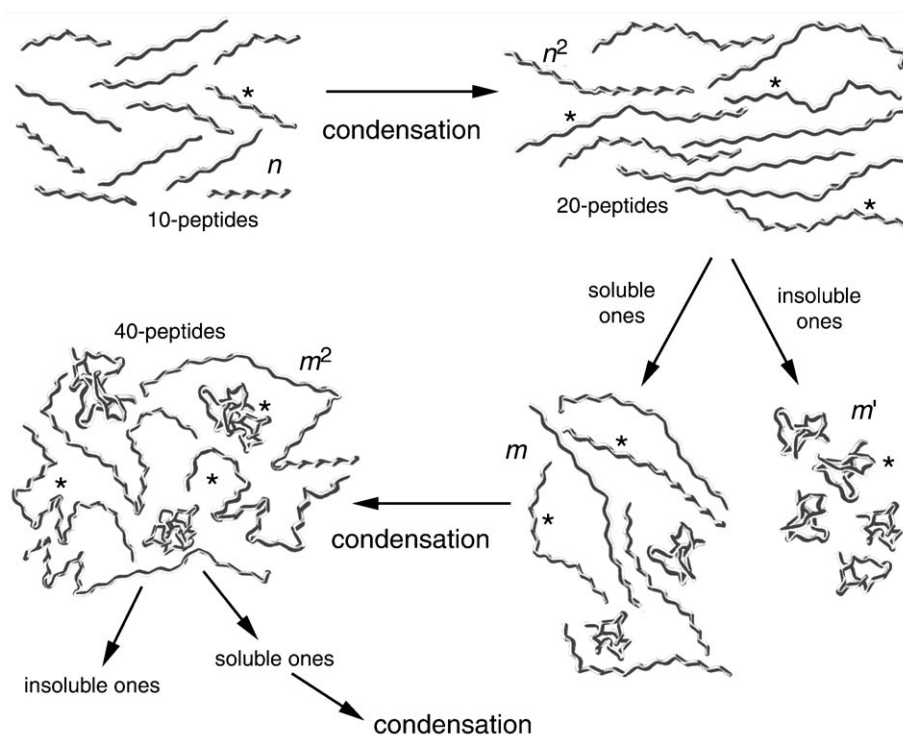


Fig. 3. The fragment condensation scheme under simulated prebiotic environmental conditions. This illustration shows how the initial library of  $n$  decapeptides may contain some compounds endowed with catalytic activity (indicated by an asterisk), and how then the ideal mutual condensation of these  $n$  decapeptides gives rise ideally to  $n^2$  20-mers, of which only  $m$  are capable of ‘*surviving*’, being soluble in  $H_2O$  under the given conditions. These, in turn, give rise to  $m^2$  40-residues-long peptides, of which a large number are insoluble under the given environmental conditions, and so on.

We have started experiments on this project, and they were already successful in obtaining a large series of relatively long NCA condensates starting from mixtures of amino acids. The problem is now to produce co-oligopeptides containing His and other catalytically important residues. As already mentioned, co-polymerization is not as easy as the polymerization of just one amino acid: starting the NCA condensation from a 1:1:1:1 mixture of four different NCA amino acids, the composition in the

copolymer may bear no relation with the initial composition of the monomer mixture. One generally obtains a library of products varying both in composition and in primary sequences, and conditions should still be worked out that permit the synthesis of a definite family of co-oligopeptides with a specific sequence (which does not have to be pre-ordered).

The next step, after having checked the reproducibility of the poly-condensation of NCAs, would be the search for the enzymatic activity in the products. And the step following that would be the attempt at fragment condensation catalyzed by such peptides.

Is it then realistic to expect that, by this method, a sizable concentration of a given long co-oligopeptide would be synthesized? The answer appears to be positive on the basis of the procedure described in the next *Section*, which reports the fragment condensation of a 44-residues-long *de novo* protein – although not based on peptide catalysts, but preliminarily only on peptide synthesis.

The whole research program is still in the initial phase, and most of the critical steps must still be worked out. It is a program of chemical synthetic biology, as we are simulating the chemistry that possibly occurred under prebiotic molecular evolution, thus reproducing a biology process. The chain elongation would proceed with a reduction of the chain candidates due to the environmental conditions, and eventually we would then obtain a sizable amount of a given, although not *a priori* programmed, co-oligo-polypeptide sequence.

*2.3. A Preliminary Experimental Implementation.* We decided to verify the validity of the theoretical scheme of the previously exposed project by utilizing, instead of catalytic peptides which are not yet at our disposal, an organic-chemistry fragment condensation based on the *Merrifield* solid-phase synthesis.

This was the work carried out in a Ph. D. program by *Salvo Chessari* under the assistance of *Richard Thomas* in my laboratory at the ETH-Zürich, and recently published [25].

First, two parent 40-residue peptides,  $P_1$  and  $P_2$ , were designed randomly but with the constraint that the relative abundance of the 20 amino acids used in their construction maintained a 1:1:1... relationship.

A matrix,  $A \cdot B$ , of 16 20-residue peptides was constructed by the systematic combination of two small libraries **A** and **B** each comprising four ten-residue peptide sequences (*Fig. 4*). The 16 20-residue sequences arrived at in this way were synthesized by the solid-phase method.

The peptide products were subjected to selection on the basis of their solubility in  $H_2O$  under well-defined conditions.

It was found that  $A_1B_2$ ,  $A_2B_2$ , and  $A_3B_2$  were completely soluble in aqueous 100 mM *Tris* buffer in the pH range of 5.2–8.6;  $A_1B_3$  and  $A_3B_3$  were insoluble, whereas  $A_2B_3$  was totally soluble, in contrast to prediction. The subsets  $(A \cdot B)_s$  that fulfilled the mentioned criterion of being soluble in  $H_2O$  were then subjected to chain elongation by combination with a further small set of 20-residue sequences, **C** (*Fig. 4*), giving rise to the new library  $C \cdot (A \cdot B)_s$  consisting of 16 peptides which are 40-residues long.

None of the latter were soluble in aqueous buffer, but two of them,  $A_1B_2C_1$  and  $A_2B_2C_1$ , turned out to be soluble in 6M guanidinium chloride (GuCl). The addition of a polar N-terminal extension to them (DDEE) resulted in the 44-residue sequences

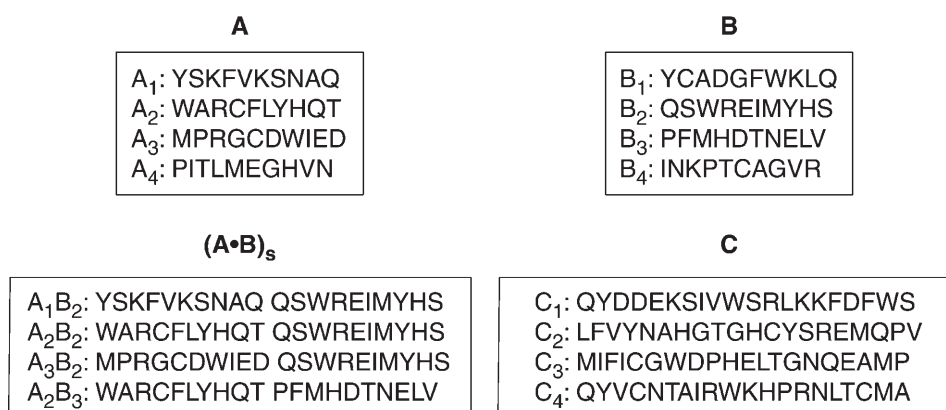


Fig. 4. The peptide sequences used for the Merrifield condensation scheme for fragment condensation

DDEE-A<sub>1</sub>B<sub>2</sub>C<sub>1</sub> and DDEE-A<sub>2</sub>B<sub>2</sub>C<sub>1</sub>. Of these two samples, only the latter was soluble in H<sub>2</sub>O. The whole sequence of this peptide is:



which was further characterized and studied.

In conclusion, then, long co-oligopeptides were obtained by the conceptual procedure illustrated previously, based namely on the fragment condensation of a larger number of smaller components, and on the elimination of most of the products due to the resultant physical properties, simply solubility in our case. Those solubility conditions, in turn, depend on the contingent environmental conditions, like pH, temperature, and solvent. Had we chosen different conditions, different peptides would have been selected out for further elongation. We are aware that, by choosing the criterion of eliminating the insoluble peptides, we may lose interesting compounds – but this is actually part of the limits of once adopted certain selection criteria.

2.4. *Spectroscopic Studies on the Folding.* It was very interesting at this point to check whether this *de novo* protein indicated above would assume a stable folding, thus reflecting even more closely the ‘birth’ of a possibly functional protein.

The far UV-CD spectrum of DDEE-A<sub>2</sub>B<sub>2</sub>C<sub>1</sub> was recorded under native and denaturing conditions as well as in the presence of 2,2,2-trifluoroethanol (TFE; data not reported here; see [25] for details). Fig. 5 shows the far UV-CD spectrum of DDEE-A<sub>2</sub>B<sub>2</sub>C<sub>1</sub> in 10 mM phosphate buffer, pH 7, at 25°. The spectrum is characteristic of a peptide that contains a high proportion of ordered structure, in particular, the presence of 49%  $\alpha$ -helix, 12%  $\beta$ -sheet, and 39% aperiodic structure.

Thermal denaturation experiments were performed over the 0–99° temperature range. The results reveal a broad, non-cooperative transition with a midpoint at *ca.* 45°. Denaturation was irreversible, and the thermally denatured form remained soluble, although the irreversibility of the process precluded any thermodynamic analysis. The

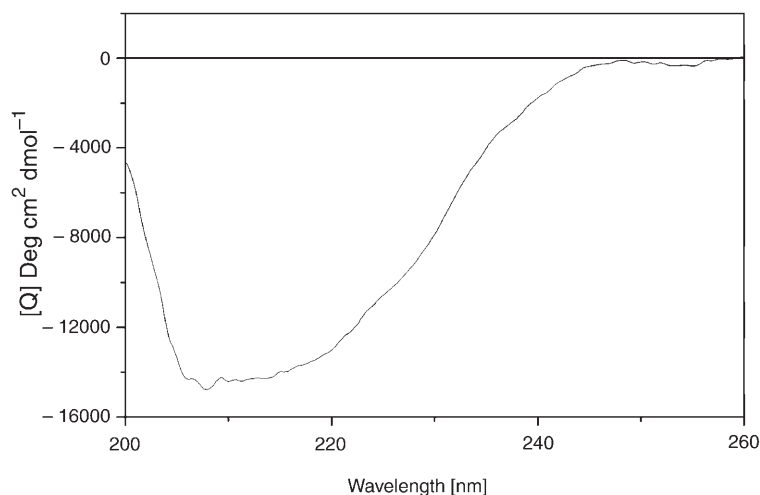


Fig. 5. CD Spectrum of  $2.5 \mu\text{M}$  DDEE- $A_2B_2C_1$  in phosphate buffer at  $25^\circ$  (see [25] for more details)

shape of the spectrum suggests that the thermally denatured form retains some structure, and that the process may reflect a conformational change rather than the complete unfolding of the peptide, a feature often observed in the thermal unfolding of small proteins.

In conclusion, one specific 44-residues-long sequence was thus obtained, which resulted from the assembly of four totally random decapeptides (plus four polar amino acids at the N-termini).

As a consequence, the 44-residues-long polypeptide sequence is to be seen as a *de novo* protein, and one that does not have any significant homologies with known sequences in the data bank for proteins of similar length. In addition, it displays a stable folding, and, in this sense, the 44-mer represents indeed the product of a model evolutionary design. Of course, *mutatis mutandis*, all what it has been stated for polypeptides can be extended to nucleic acids; the principles do not change. It should also be added that, in principle, there is no problem to obtain a few mg of such a chain with this procedure, which corresponds to an extreme large number of identical copies of such chain, according to the scheme illustrated in Fig. 3.

The choice of aqueous solubility as the decisive factor for the repetitive condensation of peptide chains was essentially arbitrary. While it offers chemical advantages, it is clear that other strategies could have been adopted, most obviously the converse, aqueous insolubility. Also, the method adopted here for the covalent chain elongation is not prebiotic. To make this more consistent with a prebiotic scenario, one might consider, for example, the fragment condensation as induced by proteolytic peptides, as mentioned in the *Introduction*.

Aside from that, the approach proposed here can be conceptually generalized to a primordial mechanism that appears capable to produce a specific macromolecular sequence from an initial oligopeptide, by a step-by-step elongation which is determined by the contingency of the environmental pressure – be pH, temperature, salinity,

solubility, aggregation, or other physical factors. This may well be a reasonable conceptual framework to conceive the etiology of specific macromolecular sequences, both in the case of polypeptides as in the case of nucleic acids.

Again, the synthetic procedure is not prebiotic, as the *Merrifield* method cannot be considered as such. As already mentioned, it is possible, however, to foresee the corresponding prebiotic synthetic scheme, based on the catalytic action of proteolytic peptides, a project that is now under scrutiny in our laboratories.

**3. The Minimal Cell Project.** – 3.1. *Premises.* The first two projects deal with the synthesis of macromolecular sequences. The one to be described now deals with the construction of synthetic – or better semi-synthetic – minimal living cell. The term semi-synthetic is meant to indicate that part of the material which is utilized, as well as the assembly procedure, is synthetic, while other parts (nucleic acid and enzymes) are of natural origin. We move more towards biology, still maintaining a good deal of a chemical synthetic approach. It is necessary from the start to make clear the limits of such a project, and to clarify the meaning of terms such as ‘*minimal*’ and ‘*living*’.

As is well-known, and as summarized in recent reviews [26–28], even the simplest unicellular organisms on Earth display a staggering complexity. *Escherichia coli* K-12 has a genome size of *ca.* 4.64 Mio (check) base pairs, and *Bacillus subtilis* of 4.2 Mio (check) base pairs, to give examples of well-known *Gram*-negative and *Gram*-positive eubacteria, respectively. The simplest known prokaryotic cell, the obligate cellular wall-less parasite *Mycoplasma genitalium*, contains 517 genes with only 470 predicted coding regions [27]. The nucleomorph chromosomes from the cryptomonad *Guillardia* has only a 551 kb genome, and, according to *Moya* and co-workers [27], *Buchnera* species have even smaller genomes that can be reduced down to 450 kb.

The question is whether such complexity is necessary for cellular life, or whether, instead, cellular life could, in principle, also be possible with a much lower number of molecular components. This proposition is relevant also for the field of the origin of life, as it does not appear reasonable to assume that life started with cells containing thousand of genes. In fact, in the field, it is generally accepted that the extant cellular complexity is the outcome of a lengthy process of evolution, starting from primordial cells that, living in a much more permissive environment, should have been genetically much simpler. How simple would a minimal cell look like, *i.e.*, a cell that contains the minimal and sufficient number of components to perform the basic functions of cellular life? Such a cell that contains the minimal and sufficient number of components to be defined as alive constitutes the notion of ‘*minimal cell*’.

And, what is ‘*living*’? The definition of life is a complex matter [9][29], and here, for the sake of generality, we will define cellular life as the capability to display a concert of three main properties: self-maintenance (metabolism), reproduction, and evolution. When these three properties are simultaneously present, we will have a full fledged cellular life. If they are not fully expressed, or if only two of them are present, we will have various forms of approximations to life, of ‘*limping life*’ [9]. These imperfect forms are also historically important, as, probably, cellular life did not start immediately with a perfect machinery. All these preliminary considerations make clear the point, that the ‘*minimal cell*’ is not a single structure, the term rather defines a family of constructs at different degrees of sophistication and complexity.

The construction of minimal cells is potentially important also from the viewpoint of biotechnology, but this is a subject which will not be dealt with here.

The notion of the minimal cell is not new, and, actually, with different emphasis and aims, the subject has been discussed several times in the literature, for example, by Woese [30], Jay and Gilbert [31], Morowitz [32], Dyson [33], Ganti [34], Szathmáry [35], and Luisi and co-workers [36–38]. Also the notion of a minimal RNA cell containing only a couple of RNA genes in a self-reproducing vesicle has been presented [39].

3.2. *The Minimal Genome.* If one accepts the idea that cells can be living with a lower complexity, one should first consider the question of the minimal genome.

This question has been considered by several authors, for example, Mushegian and Koonin [40–42], Shimkets [43], Kolisnychenko *et al.* [44], Luisi *et al.* [26], Gil *et al.* [45], Islas *et al.* [28], and Pohorille and Deamer [46].

Mushegian and Koonin [40] calculated an inventory of 256 genes that represents the amount of DNA required to sustain a modern type of minimal cell under permissible conditions. Craig Venter and his group, using a ‘knocking down’ approach, arrived in the case of *M. genitalium* bacterium to the conclusion that *ca.* 265 to 350 genes are essential under laboratory growth conditions [7].

Andres Moya and his group in Valencia arrive at the smaller number of 206 genes in the case of *Buchnera* and other organisms [27][45]. The number of 206 genes as minimal genome represents, on one hand, a considerable simplification. On the other hand, it still corresponds to a formidable complexity. This arises again the question, whether and how can one go further down.

Further speculations are in fact presented in the literature. For example, taking the case of *Mycoplasma genitalium*, one can strip its genome by several enzymes and factors, arriving at *ca.* 150 genes [26]. This figure, by a process of more drastic assumptions, can be further reduced to 40–50 by eliminating the ribosomal proteins and reducing the specificity of the enzymatic processes. Of course, there is not a way to demonstrate that these simplified cells would work, but this is anyway an interesting and useful theoretical procedure in order to conceive how to start the corresponding experimental work of construction.

3.3. *Experimental Approaches to the Minimal Cell.* Having clarified, to some extent, the nature of the minimal genome, one might try to have it inserted into a cell, to see whether it works. But it is not easy to prepare the genome of *Buchnera* or any other micro-organism with any desired number of genes.

The alternative approach which has been considered in my and other groups is illustrated schematically in Fig. 6.

The idea is to construct a minimal cell by inserting in a compartment a calibrated ensemble of genes and/or enzymes, increasing step by step the complexity of this added ensemble, until the construct begins to display forms of cellular life. In my group, this work is carried out by Yutetsu Kuruma, Giovanni Murtas, and Pasquale Stano.

Liposomes have been chosen as membrane compartments, since they, with their lipid bilayer, are considered the best models for the shell of biological membranes, and the procedure to insert biopolymers inside liposomes is already well-established. Another good reason to use vesicles or liposomes is due to the fact that conditions have been described under which vesicles are capable of self-reproduction [47–49], namely,

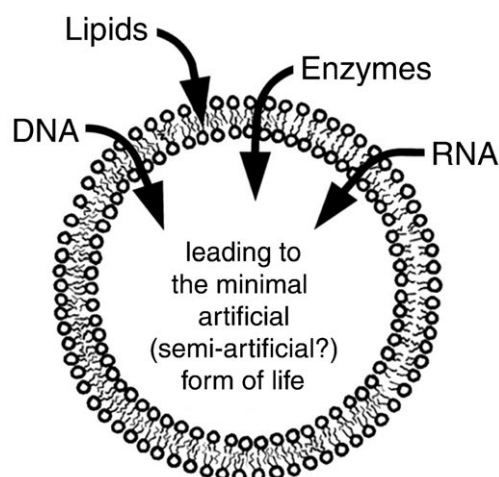


Fig. 6. *The semi-synthetic approach to the construction of the minimal cell*

of increasing auto-catalytically their population number, which may model the behavior of living cells.

The constructs obtained according to the procedure of *Fig. 6* are defined as semi-synthetic, since, as already mentioned, the compartment and the technical procedure are synthetic, but the materials (enzymes and nucleic acids) are taken from nature. Note also that this is not a program about the origin of life, as one starts from extant and mature macromolecules.

To arrive, in this way, at a minimal cellular life is a complex enterprise, and it is useful to divide up the 'road map' to the minimal cell in different milestones with increasing complexity.

The first one, which is already under control in several laboratories, is to carry out and optimize complex enzymatic reactions in the interior of liposomes, such as the polymerase chain reaction, the biosynthesis of RNA and DNA, the condensation of amino acids, *etc.*

All these reactions which have been carried out in liposomes have been reviewed a few times in the last years [36][50] and will not be repeated here. It may be interesting to recall here only two of them.

Most probably, the very first attempt to carry out biological reactions inside liposomes with the aim of creating a minimal cell is the work by *Schmidli et al.* [51]. The idea here was to enzymatically synthesize lecithin inside lecithin-liposomes: in this way, there would be a self-reproduction caused by an internal reaction, in keeping also with the basic ideas of autopoiesis [52–54].

The other and only other example that I like to mention here is one based on the above mentioned self-reproduction of oleate vesicles. A suggestive example of core and shell reproduction was in fact provided in 1995 by *Oberhozer et al.* with the use of  $Q\beta$  replicase [55].

As illustrated in the *Fig. 7*, while the enzyme was replicating RNA inside the vesicles, the oleate vesicles were multiplying by their own accord. At first sight, this

appears already a good approximation of the living cell: there is a simple metabolism inside it, the system is replicating, and potentially capable of evolution, since there is a copying machinery of RNA that can give rise to mistakes/mutations. However, the enzyme and the RNA molecules are not reproduced from inside the system, and, after a few generations, most of the new vesicles will not be capable of further reproduction: in fact, for statistical reasons, they will not contain all original system's components, most of them will be empty or containing only one macromolecular component. The system will undergo what we call 'death by dilution' [9].

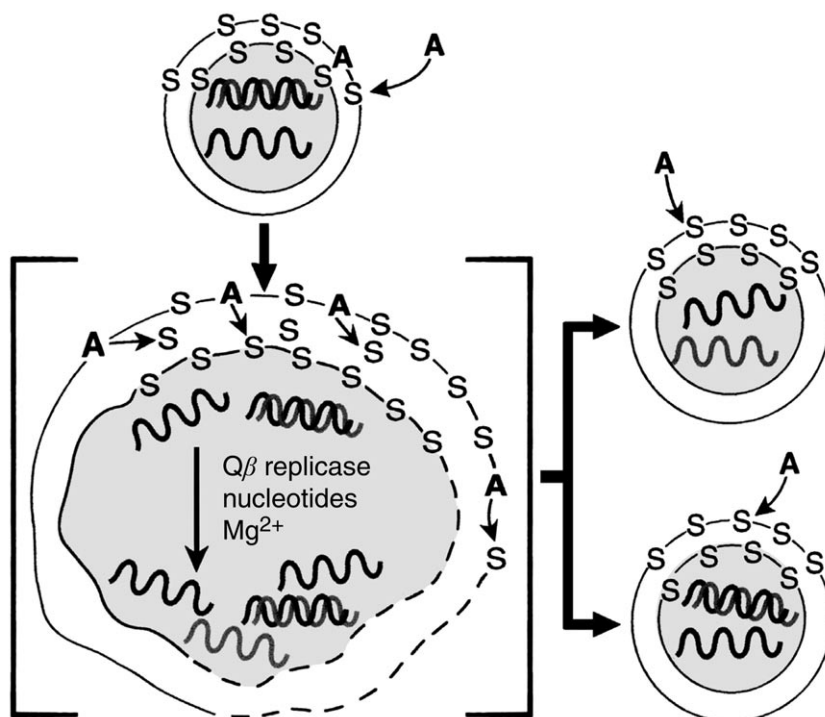


Fig. 7. A core-and-shell reproduction with oleate vesicles and the  $Q\beta$  replicase system. Both the vesicle shells and the vesicle content are replicating, although in an uncoupled way (see [55]).

**3.4. Semi-Synthetic Cells as Examples of Synthetic Biology.** The next step in the road map to the minimal cell is the encapsulation, in the water pool of vesicles, of all components which are necessary to express proteins. Mostly, the green fluorescence protein (GFP) has been expressed until now, as it is easier to detect it. Among the research groups which are active in this field, one should mention those one by Yomo and co-workers [56], as well as Nomura *et al.* [57] in Japan, Noireaux and Libchaber in the States [58], our group first in Zurich and now in Rome [36–38][59].

In all these cases, the entire ribosomal machinery has been incorporated into the liposomes. Generally, commercial kits for protein expression have been utilized, for which the composition and the relative concentrations of components are not known.



We are then dealing with ‘*black boxes*’, which contain most probably a few hundreds of enzymes.

One significant progress in the field is the description by *Ueda* and co-workers [60] of the so-called PURE SYSTEM. This is a system consisting of 37 enzymes (plus the ribosomes and tRNAs), which is capable of expressing proteins *in vitro*. Recently, *Murtas* and *Kuruma* were capable of inserting the whole system in liposomes, and expressing GFP and two proteins related to the lipid metabolism [61]. This is already a considerable record in the field of the minimal cell, as we are dealing with a system capable of full protein expression and having only 37 enzymes for that (although the necessary number of genes is larger than that).

Protein expression is only part of the life of a cell. To have a satisfactorily minimal cell, one should reach the point at which the minimal cell is also able to self-reproduce; and we are working on that. The idea is to enrich the 37 components of the PURE SYSTEM with the minimal number of additional enzyme which permit the synthesis of the ribosomes themselves. In another direction of work, we will try to simplify the structure of the ribosomes, in order to eliminate most of the proteins.

**4. Concluding Remarks.** – We have presented three projects that can be classified in the field of chemical synthetic biology, where the characterization of the single chemical constituents is still the major or one of the major features. In the NBPs project, the synthetic-biology products are single proteins that do not exist in nature, and, actually, the same holds for the second project presented here, where the synthesis, however, obeys a more classical organic-synthesis pattern. The project on the minimal cell is certainly the most biological of the three; however, the procedure is based on the preparation and physico-chemical characterization of vesicles of given dimensions, and the insertion of macromolecular species of known composition and concentration – a typical chemistry approach. Of these three projects, the first and the third one can be considered outcomes of modern techniques in molecular biology – in the sense that they were not technically possible ten or twenty years ago. By contrast, the second one, dealing with the prebiotic synthesis of specific macromolecular sequences, does not require any advanced modern technical skill, and one may wonder why it has not been attempted or implemented long ago. Interestingly, the minimal cell project can be seen as a concrete example of bridging systems and synthetic biology, being focussed on a constructive approach of a cellular system.

It is important for me to mention that this experimental work has a philosophical counterpart, in the sense that it bears with fundamental questions in the field of the origin of life. Thus, the NBP project is germane to the question of the selection of our extant proteins, and, therefore, connects to the controversy between determinism and contingency in the things of nature; and so is for the second project. The third one, on the minimal cell, has directly to do with questions such as ‘*what is life*’ and with the question, whether life is an emergent property arising from non-living components.

This interface between chemistry and philosophy is arising in a special way when chemistry moves towards biology, and, of course, synthetic biology is the most proper medium for this merging.

The Author thanks his co-workers, first *Pasquale Stano*, and then *Cristiano Chiarabelli*, *Davide de Lucrezia*, and *Peter Strazewski* of the University of Lyon, for the very useful discussions and modification proposals of the manuscript.

## REFERENCES

- [1] G. Church, *Nature* **2005**, *438*, 423; and all articles in this special issue.
- [2] D. Ferber, *Science* **2004**, *303*, 157; and all articles in this special *Science* issue.
- [3] M. Bolli, R. Micura, A. Eschenmoser, *Chem Biol.* **1997**, *4*, 309; see also M. Bolli, R. Micura, S. Pitsch, A. Eschenmoser, *Helv. Chim. Acta* **1997**, *80*, 1901.
- [4] S. A. Benner, A. M. Sismour, *Nature Rev. Genet.* **2005**, *6*, 524.
- [5] N. Doi, K. Kakukawa, Y. Oishi, H. Yanagawa, *Protein Eng. Design Selection* **2005**, *18*, 279.
- [6] S. Akanuma, T. Kigawa, S. Yojoyama, *Proc. Natl. Acad. Sci U.S.A.* **2002**, *99*, 13549.
- [7] C. A. Hutchinson III, S. N. Peterson, S. R. Gill, R. T. Cline, O. White, C. M. Fraser, H. O. Smith, J. C. Venter, *Science* **1999**, *286*, 2165; C. Zimmer, *Science* **2003**, *299*, 1006.
- [8] C. de Duve, 'Life Evolving: Molecules, Mind and Meaning', Oxford University Press, New York, 2002.
- [9] P. L. Luisi, 'The Emergence of Life: From Chemical Origins to Synthetic Biology', Cambridge University Press, Cambridge, 2006.
- [10] W. Paley, 'Natural Theology: or Evidences of the Existence and Attributes of the Deity, Collected from the Appearances of Nature', 1802; reprinted 1972 by St. Thomas Press, Houston, Texas, p. 3.
- [11] J. D. Barrow, F. J. Tipler, *Nature* **1988**, *331*, 31; J. D. Barrow, *Ann. N.Y. Acad. Sci.* **2001**, *950*, 139.
- [12] J. D. Barrow, F. J. Tipler, 'The Anthropic Cosmological Principle', Oxford University Press, Oxford 1996.
- [13] S. J. Gould, 'Wonderful Life', Penguin Books, London, 1989.
- [14] J. Monod, 'Chance and Necessity', Knopf, New York, 1971.
- [15] a) C. Chiarabelli, J. W. Vrijbloed, R. M. Thomas, P. L. Luisi, *Chem. Biodiv.* **2006**, *3*, 827; b) C. Chiarabelli, J. W. Vrijbloed, D. De Lucrezia, R. M. Thomas, P. Stano, F. Polticelli, T. Ottone, E. Papa, P. L. Luisi, *Chem. Biodiv.* **2006**, *3*, 840; c) D. De Lucrezia, M. Franchi, C. Chiarabelli, E. Gallori, P. L. Luisi, *Chem. Biodiv.* **2006**, *3*, 860; d) D. De Lucrezia, M. Franchi, C. Chiarabelli, E. Gallori, P. L. Luisi, *Chem. Biodiv.* **2006**, *3*, 869.
- [16] a) L. E. Orgel, *Orig. Life Evol. Biosphere* **1998**, *28*, 227; b) M. Paechthorowitz, F. R. Eirich, *Orig. Life Evol. Biosphere* **1988**, *18*, 359; c) H. Tsukahara, E.-i. Imai, H. Honda, K. Hatori, K. Matsuno, *Orig. Life Evol. Biosphere* **2002**, *32*, 13.
- [17] J. P. Ferris, *Orig. Life Evol. Biosphere* **2002**, *32*, 311.
- [18] S. W. Fox, K. Dose, 'Molecular Evolution and the Origin of Life', Marcel Dekker, New York, 1977.
- [19] J. Taillades, H. Cottet, L. Garrel, I. Beuzelin, L. Boiteau, H. Choukroun, A. Commeyras, *J. Mol. Evol.* **1999**, *48*, 638.
- [20] A. Commeyras, L. Boiteau, O. Vandenabeele-Trambouze, F. Selsis, in 'Lectures in Astrobiology' Vol. 1, Part 2, Eds. M. Gargaud, B. Barbier, H. Martin, and J. Reisse, Springer-Verlag, Berlin, 2004, p. 517–542.
- [21] Y. Li, Y. Zhao, S. Hatfield, R. Wan, Q. Zhu, X. Li, M. McMills, Y. Ma, J. Li, K. L. Brown, C. He, F. Liu, X. Chen, *Bioorg. Med. Chem.* **2000**, *8*, 2675.
- [22] K. Plankensteiner, A. Righi, B. M. Rode, *Orig. Life Evol. Biosphere* **2002**, *32*, 225.
- [23] H.-D. Jakubke, P. Kuhl, A. Könnicke, *Angew. Chem., Int. Ed.* **1985**, *24*, 85; H.-D. Jakubke, in: 'The Peptides: Analysis, Synthesis, Biology', Eds. S. Udenfried and J. Meienhofer, Academic Press: New York, 1987, Vol. 9, chapter 3; H.-D. Jakubke, in: 'Enzyme Catalysis in Organic Synthesis', Eds. K. Drauz, and H. Waldmann, VCH, Weinheim, 1995, Vol. I, p. 431; H.-D. Jakubke, U. Eichhorn, M. Hänsler, D. Ullmann, *Biol. Chem.* **1996**, *377*, 455.
- [24] R. Jost, E. Brambilla, J. C. Monti, P. L. Luisi, *Helv. Chim. Acta* **1980**, *63*, 375; A. Pellegrini, P. L. Luisi, *Biopolymers*, **1978**, *17*, 2573.
- [25] S. Chessari, R. Thomas, F. Polticelli, P. L. Luisi, *Chem. Biodiv.* **2006**, *3*, 1202.
- [26] P. L. Luisi, T. Oberholzer, A. Lazcano, *Helv. Chim. Acta* **2002**, *85*, 1759.

- [27] R. Gil, B. Sabater-Munoz, A. Latorre, F. J. Silva, A. Moya, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4454.
- [28] S. Islas, A. Becerra, P. L. Luisi, A. Lazcano, *Orig. Life Evol. Biosphere* **2004**, *34*, 243.
- [29] P. L. Luisi, *Orig. Life Evol. Biosphere* **1998**, *28*, 613.
- [30] C. R. Woese, 'The Primary Lines of Descent and the Universal Ancestor', in 'Evolution: From Molecules to Man', Ed. D. S. Bendall, Cambridge University Press, Cambridge, 1983, p. 209–233.
- [31] D. Jay, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 1978.
- [32] J. Morowitz, 'Beginnings of Cellular Life', Yale University Press, 1992.
- [33] F. J. Dyson, *J. Mol. Evol.* **1982**, *18*, 344.
- [34] T. Ganti, 'The Principles of Life', Oxford University Press, 2003.
- [35] E. Szathmáry, *Nature* **2005**, *433*, 469.
- [36] P. L. Luisi, *The Anatomical Record* **2002**, *268*, 208.
- [37] P. L. Luisi, F. Ferri, P. Stano, *Naturwissenschaften* **2006**, *93*, 1.
- [38] T. Oberholzer, P. L. Luisi, *J. Biol. Phys.* **2002**, *28*, 733.
- [39] J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **2001**, *409*, 387.
- [40] A. Mushegian, E. Koonin, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10268.
- [41] E. V. Koonin, *Nat. Rev. Microbiol.* **2003**, *1*, 127; E. V. Koonin, *Annu. Rev. Genomics Human Genet.* **2000**, *1*, 99.
- [42] A. Mushegian, *Curr. Opin. Genet. Develop.* **1999**, *9*, 709.
- [43] L. J. Shimkets, 'Structure and Sizes of Genomes of the Archaea and Bacteria', in: 'Bacterial Genom', Eds. F. J. De Bruijn, J. R. Lupskin, and G. M. Weinstock, 1998.
- [44] V. Kolisnychenko, G. Plunkett III, C. D. Herring, T. Fehér, J. Pósfai, F. R. Blattner, G. Pósfai, *Genet. Res.* **2002**, *12*, 640.
- [45] R. Gil, F. J. Silva, J. Peretó, A. Moya, *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 518.
- [46] A. Pohorille, D. Deamer, *Trends Biotechnol.* **2002**, *20*, 123.
- [47] P. A. Bachmann, P. L. Luisi, J. Lang, *Nature* **1992**, *357*, 57.
- [48] P. Walde, R. Wick, M. Fresta, A. Mangone, P. L. Luisi, *J. Am. Chem. Soc.* **1994**, *116*, 11649.
- [49] R. Wick, P. Walde, P. L. Luisi, *J. Am. Chem. Soc.* **1995**, *117*, 1435.
- [50] P. L. Luisi, T. Oberholzer, 'Origin of Life on Earth: Molecular Biology in Liposomes as an Approach to the Minimal Cell', in 'The Bridge between the Big Bang and Biology', Ed. F. Giovannelli, CNR Press, 2001, p. 345–355.
- [51] P. K. Schmidli, P. Schurtenberger, P. L. Luisi, *J. Am. Chem. Soc.* **1991**, *113*, 8127.
- [52] F. Varela, H. R. Maturana, R. B. Uribe, *Biosystems* **1974**, *5*, 187.
- [53] F. Varela, H. Maturana, 'The Tree of Knowledge', Shambala, Boston, 1998.
- [54] P. L. Luisi, *Naturwissenschaften* **2003**, *90*, 49.
- [55] T. Oberholzer, R. Wick, P. L. Luisi, C. K. Biebricher, *Biochem. Biophys. Res. Commun.* **1995**, *207*, 250.
- [56] K. Ishikawa, K. Sato, Y. Shima, I. Urabe, T. Yomo, *FEBS Letters* **2004**, *576*, 387; W. Yu, K. Sato, M. Wakabayashi, T. Nakatshi, E. P. Ko-Mitamura, Y. Shima, I. Urabe, T. Yomo, *J. Biosci. Bioeng.* **2001**, *92*, 590; T. Sunami, K. Sato, T. Matsuura, K. Tsukada, I. Urabe, T. Yomo, *Anal. Biochem.* **2006**, *357*, 128.
- [57] S. M. Nomura, K. Tsumoto, T. Hamada, K. Akiyoshi, Y. Nakatani, K. Yoshikawa, *ChemBioChem* **2003**, *4*, 1172.
- [58] V. Noireaux, A. Libchaber, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17669.
- [59] A. V. Pietrini, P. L. Luisi, *ChemBioChem* **2004**, *5*, 1055; D. Fiorimondo, P. Stano, P. L. Luisi, in preparation.
- [60] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat. Biotechnol.* **2001**, *19*, 751.
- [61] G. Murtas, Y. Kuruma, P. L. Luisi, *BMC Chem. Biol.*, in press.

Received November 16, 2006