

A UNIVERSAL DEFINITION OF LIFE: AUTONOMY AND OPEN-ENDED EVOLUTION

KEPA RUIZ-MIRAZO^{1,2}, JULI PERETÓ³ and ALVARO MORENO^{1,4*}

¹Centre of Astrobiology (INTA/CSIC), Madrid, Spain; ²Department of Fundamental Sciences, Mondragón University, Basque Country, Spain; ³Department of Biochemistry and Molecular Biology, University of València, Spain; ⁴Department of Logic and Philosophy of Science, University of the Basque Country, San Sebastian-Donostia, Spain
(* author for correspondence, e-mail: ylpmobeas@sf.ehu.es)

(Received 1 October 2002; accepted in revised form 12 December 2002)

Abstract. Life is a complex phenomenon that not only requires individual self-producing and self-sustaining systems but also a historical-collective organization of those individual systems, which brings about characteristic evolutionary dynamics. On these lines, we propose to define universally living beings as autonomous systems with open-ended evolution capacities, and we claim that all such systems must have a semi-permeable active boundary (membrane), an energy transduction apparatus (set of energy currencies) and, at least, two types of functionally interdependent macromolecular components (catalysts and records). The latter is required to articulate a 'phenotype-genotype' decoupling that leads to a scenario where the global network of autonomous systems allows for an open-ended increase in the complexity of the individual agents. Thus, the basic-individual organization of biological systems depends critically on being instructed by patterns (informational records) whose generation and reliable transmission cannot be explained but take into account the complete historical network of relationships among those systems. We conclude that a proper definition of life should consider both levels, individual and collective: living systems cannot be fully constituted without being part of the evolutionary process of a whole ecosystem. Finally, we also discuss a few practical implications of the definition for different programs of research.

Keywords: artificial life, astrobiology, autonomous agents, definition of life, generalization of biology, genotype-phenotype decoupling, open-ended evolution, origin of life, origins of (genetic) information

1. Introduction

Definitions of life are highly controversial. And it is not just a question of lack of consensus among the different proposals formulated so far. Some authors are very skeptical about the actual possibility of grasping 'in any scientifically relevant language' such a complex and multifarious phenomenon. Others think that we have to wait until biological theory(ies) become more rigorous, more encompassing and meaningful. And some others consider that it is not worth undertaking the challenge since, even if we could obtain a proper definition of life, it would still be a rather conventional one and would probably have little influence on the development of specific research programs in biology.



Origins of Life and Evolution of the Biosphere **34**: 323–346, 2004.
© 2004 Kluwer Academic Publishers. Printed in the Netherlands.

The living phenomenology shows, indeed, many different sides (that appear at various levels of organization) and it is not easy to capture all of them in a single conceptual scheme. This is made even more difficult by life's ability to diversify and explore its own limits (always producing border-line cases, exceptions to the rule, ...). Last century's impressive advances in molecular biology have revealed a great underlying biochemical unity of all living forms, but it is not clear to what extent this is the result of contingency or of real necessity: i.e., whether that unity can serve to extract general biological principles or just derives from having a universal common ancestor of all terrestrial life. In addition, since the problem of the origin of life is also far from being solved, it is not at all obvious how those 'biological principles' would relate to the general laws of physics and chemistry, i.e., if they would be subject to an eventual reduction to the latter, or should have their own 'status' (with their own explanatory power, degree of abstraction, etc.) as scientific laws.

However, despite these and other difficulties that we could think of, there are also good arguments that now is a suitable time to tackle the question (like some authors are, in fact, doing, see e.g. Pályi *et al.* (2002)). Certainly, no general full-fledged theory of biology is available yet (as Cleland and Chyba (2002) highlight in order to support their skeptical view), but the insightful research carried out during the last decades in areas such as bioenergetics, enzymology or genetics provide us with a body of knowledge which is deep and wide enough to try structuring it around some fundamental 'tentatively universal' concepts. This effort, in itself, can be very helpful for the development of a general, and better formalized, biology. The lack of success of previous attempts at defining life (in the sense that they have not led to a well-established consensus) should not discourage us, especially after noticing that some confusion might still exist in the field (see the article on the subject recently published in *Science* by Koshland (2002), as well as our critical remarks below). And neither should the skeptical minds of those who do not regard the task of defining life as a possible or useful one for biology. Their claims, based on the continuity between physico-chemical and biological phenomena, or on the irreducibility/reducibility of biology to physics and chemistry, are not conclusive.

In this article we suggest a concrete definition of life (of minimal and universal life), with the aim of opening it to discussion. Our proposal is meant not only to provide some conceptual clarification on this particular issue, but to contribute to guide other lines of research, like specific programs in the fields of origin of life, astrobiology or artificial life. The first step, anyway, is to determine what kind of definition serves the purpose and, then, justify in what sense the proposal here suggested means an advance with regard to previous attempts.

2. Requirements a Definition of Life Should Fulfill

In general terms, definitions can be made with two main different purposes: (i) to demarcate or classify a certain type of phenomenon, and (ii) to make manifest – and, perhaps, even explain – the fundamental nature of that type of phenomenon. The first purpose normally leads to *descriptive* definitions that consist in a set of properties (typically, a ‘check-out’ list) containing all what is required to determine if a phenomenon belongs to a particular kind or not, whereas the second involves a completely different way of formulating the question: *essentialist* definitions characterize a given phenomenon in terms of its most basic dynamic mechanisms and organization.

Several lists of properties have been suggested in the literature to discern ‘the inert’ from ‘the living’, e.g. self-organization, growth, development, functionality, metabolism, adaptability, agency, reproduction, inheritance or susceptibility to death. See, among others, Mayr (1982), Farmer and Belin (1992), or the so-called ‘seven pillars’ of de Duve (1991). However, since these catalogues do not provide a hierarchy or an account through which the chosen properties can be related to each other, it is hard to tell whether any of the proposals includes just the necessary and sufficient ones (if there are no properties deriving from others – i.e. the set would be redundant – or if some additional fundamental property ought to be introduced – i.e. there is something important missing in it). Besides, lists do not offer any hint to clarify the source or process of integration of the system/phenomenon under analysis. This is crucial because a definition of life that is expected to be truly universal must be built from general principles, some of which should stem from physics and chemistry. In other words, the definition has to include primitive concepts that help to bridge the gap between physico-chemical and biological phenomenology. This was already highlighted by Oparin (1961), who claimed that the problem of defining life is tightly intertwined with the problem of its origin.

Therefore, we must look for an essentialist type of definition that, at the same time, is ‘genealogical’, in the sense that it explains – or at least throws some light on – the process that leads to the constitution of the phenomenon, starting from well-determined conditions. In this way, it should offer a natural framework for generalization (i.e., for selecting which are the universal features of all possible life). More schematically, the set of requirements that a definition of life should meet in order to be of use for the development of present biology – and other related fields of research – can be condensed in the following points (by and large equivalent to Emmeche’s (1998)):

The definition should:

- (a) be fully coherent with current knowledge in biology, chemistry and physics;
- (b) avoid redundancies and be self-consistent;

- (c) possess conceptual elegance and deep explanatory power (i.e., it must provide a better understanding of the nature of life, guiding our search into its origins and its subsequent maintenance and development);
- (d) be universal (in the sense that it must discriminate the necessary from the contingent features of life, selecting just the former);
- (e) be minimal but specific enough (i.e., it should include just those elements that are common to all forms of life – not being, in principle, restricted to life on Earth – and, at the same time, it must put forward a clear operational criterion to tell the living from the inert, clarifying border-line cases, contributing to determine biomarkers, etc.).

According to these general points, we can assess the validity or usefulness of different proposals. Lists, for instance, show difficulties to fulfill 'c', and typically also 'b' (and, thereafter, 'd' and 'e'). Vitalist approaches neglect 'a'. So those are already 'out of the game'. Let us briefly check out a recent example: Koshland's proposal (2002). This author offers a definition that is not intended to be a mere list, but it ends up exhibiting very similar weaknesses. His suggestion of a set of seven 'principles' or 'pillars' of life (program, improvisation, compartmentalization, energy, regeneration, adaptability and seclusion¹) is not really satisfactory: it not only lacks elegance and explanatory power, but is clearly redundant (several of Koshland's 'pillars' can be included in the concept of 'autopoiesis', see below and also contrast with Table I in Margulis (1990)).

Obviously, we can find other options in the literature, which fit better into the group of essentialist definitions. To be highlighted here is Emmeche's *bio-semiotic* definition (1998): 'life is the functional interpretation of signs in self-organizing material 'code-systems' that construct their own 'umwelts' '. This proposal is certainly self-consistent and conceptually elegant (even more so if one looks into the bio-semiotic theory), but has an important drawback: it assumes 'signs' (i.e., information) as a primitive natural kind, when physics and chemistry do not (the latter just accept: time, matter, energy, charge, fields, ...). Therefore, although it does not directly contradict physico-chemical knowledge, it does not come to terms with it, and requirement 'a' is not strictly fulfilled.

It is also worth mentioning here Shapiro and Feinberg's definition (1990), according to which life is the activity of a 'biosphere', i.e., the activity of 'a highly ordered system of matter and energy characterized by complex cycles that maintain or gradually increase the order of the system through an exchange of energy with its environment'. This proposal is formulated in very broad terms and with a well-intended purpose of bringing biology closer to physics. However, the apparent advantage of this formulation (particularly in relation to requirements 'a' and 'd') turns out to be a shortcoming: where is the specificity of the living phenomenon? In this scheme, where can we look for properties like self-production, re-production, adaptation, heredity, selective evolution ..., which are so characteristic of biological systems? The motivation to establish a link between the inert

and the living domains should not lead us to forget all that we actually know about biology (or chemistry!). There are many ways of maintaining or increasing the order of a system whose dynamics takes place in far from equilibrium conditions (self-organizing phenomena), but living systems have developed their own, and that is why they are so distinct from anything else. Thus, the definition above is not suitable since it should, at least, give an indication of where to search in this direction (i.e., requirement 'e' is not adequately fulfilled).

The issue is quite tricky: we cannot expect that all biologically relevant items are touched or made manifest in a brief statement like a definition. But, nevertheless, a good definition must be well-balanced, containing the key conceptual tools to develop a theory or scheme around it that is coherent and provides specific enough hints to establish a natural connection between the physico-chemical and biological realms. In the following section we will review two examples that get closer to this ideal. Even if they are 'subject to improvement', we can say that, by and large, they meet all the standards above. And, in any case, the discussion on these two definitions will be very helpful to introduce our own later.

3. Brief Review of Two Significant Attempts

The main message to be gathered from the last section is that a proper definition of life should not be merely descriptive but 'essentialist-generational'. In particular, we must search for definitions that contain high explanatory power/deepness, i.e., that are supported by or tightly linked to a well-founded conceptual framework about biological phenomenology. On these lines, two major proposals will be selected from the literature.

3.1. STANDARD DEFINITION

'Life is a system capable of evolution by natural selection' (Sagan, 1970) or, in a more elaborate and precise version: 'Life is a self-sustained chemical system capable of undergoing Darwinian evolution' (NASA's 'working definition'; Joyce, 1994; see also Luisi, 1998).

The underlying conception is very close to the standard view that supports the neo-Darwinian paradigm in biology, where the stress is put on the evolutionary dynamics of biological systems. According to this view, the key properties that a system must show in order to evolve through natural selection are 'reproduction, variability and inheritance' (Lewontin, 1970; Maynard Smith, 1986). It follows, then, that any system with those properties, including a population of replicating molecules, could readily fulfill the definition.

Some authors, like Wicken (1987), Brooks and Wiley (1988), Luisi (1998) or Kauffman (2000), are critical of this perspective, arguing that a proper definition of living being must take into account the characteristic way in which the

components of such a system get organized as a coherent whole. The second version of the definition might seem less vulnerable to this criticism, since it introduces the idea of 'self-maintenance' (apparently linked to that of metabolism, see below). However, the core idea stays the same. In Joyce's own terms (1994, p. xi) while 'the notion of Darwinian evolution subsumes the processes of self-reproduction, material continuity over an historical lineage, genetic variation, and natural selection', 'self-maintenance' 'refers to the fact that living systems contain all the genetic information necessary for their own constant production (i.e., metabolism)'. The geneticist bias is rather obvious: metabolism here seems to be the result of acquiring a complete enough pool of genes so as to achieve the constant production of a system. As Luisi (1998) emphasizes, this definition was created from a conception of life that is fundamentally molecular, in tune with a general research program that looks into the roots of Darwinian evolution in the context of populations of replicating molecules undergoing some selective dynamics (e.g., models of 'quasi-species', 'hypercycles', etc.; see Eigen and Schuster, 1979; Eigen, 1992).

Nevertheless, as the next definition will show, the notion of metabolism is much richer than that. It involves a material self-(re-)producing organization, functionally integrated (i.e., not reducible to the properties of its molecular components), which is not really grasped by the term 'self-maintenance' (specially if this is understood in a weak sense; see Moreno and Ruiz-Mirazo, 1999; Ruiz-Mirazo and Moreno, 2000). In summary, the basic problem of this definition is that there is no proper characterization of the type of material organization that would allow the beginning of a process of Darwinian evolution (precisely, some sort of pre-genetic metabolic organization).

3.2. AUTOPOIETIC DEFINITION

'An autopoietic system – the minimal living organization – is defined as a network of processes of production (synthesis and destruction) of components such that these components: (i) continuously regenerate and realize the network that produces them, and (ii) constitute the system as a distinguishable unit in the domain in which they exist' (Varela, 1994; originally Maturana and Varela, 1973).

This definition, despite its high level of abstraction, is much more explicit than the previous one with regard to capturing the basic organization that constitutes a minimal living being, its metabolism. The main idea put forward in it is that life should be defined from the perspective of individual organisms. A living being is conceived as a recursive web of component production and transformation, continuously self-generated and self-regenerated, which produces a physical boundary that, in turn, is a necessary condition for the maintenance of the web. All this conveys an organizational 'autos' that goes beyond what is usually considered as a phenomenon of 'self-organization'. This 'autos' involves a group of interrelated processes that gets organized according to a global operational logic: the system is,

somehow, 'closed up in itself', in the sense that it creates its own boundaries and follows its own circular dynamics of self-production.

Yet, the autopoietic definition presents two main problems.

- (1) On the one hand, it is excessively abstract. Maturana and Varela's original purpose was to define a minimal biological system in as general terms as possible. However, they go too far in that direction, since they offer a conception of the living which is too detached from the physical – material and energetic – requirements that are crucial for its actual implementation (as Fleischaker (1988) already highlighted). This leads to a characterization of the 'minimal organizational logic' of biological systems in which physics and chemistry (thermodynamics in particular) have nothing to say, and in which the interactive-agential relationship between system and environment is disregarded (or, at least, it is regarded as secondary). However, the introduction of material-thermodynamic aspects involves a totally different 'organizational logic' of self-production (Moreno and Ruiz-Mirazo, 1999; Kauffman, 2000), highlighting the relationship with the environment as one of the key ingredients to understand how a minimal self-constructing chemical system operates (Ruiz-Mirazo and Moreno, 2000)².
- (2) On the other hand, even if the autopoietic criteria were modified so as to include some additional ingredients related to the previous point, it would not suffice to characterize 'the living' in a complete way. The ability to produce diversity and growth in complexity as a result of a selective evolutionary process, which is another fundamental feature of biological systems, is not at reach for autopoietic systems. In principle, an autopoietic system is capable of reproduction (through auto-catalytic growth and division), of adapting to external disturbances (organizational homeostasis) and even of modifying its type of organization – its identity, in a broad sense – (through the accumulation of structural changes), but it is not capable of initiating a process of Darwinian evolution because, as such, it does not have the genetic mechanisms required to do so. Thus, the main problem of this definition, the characteristic evolutionary capacity of living beings, is not taken into account.

Quite interestingly, these two definitions (standard and autopoietic) cover the two fundamental aspects of biological phenomenology. As Maynard Smith (1986) already pointed out, life appears both as a collective population of self-reproducing hereditary systems (life as evolution) and as individual self-maintaining dissipative units (life as metabolism). The problem of both definitions is that each focuses on just one of those two dimensions: neither of them works out properly – i.e., in a well-balanced way – the tension between the individual-metabolic and the collective-ecological sides of the phenomenon. Nevertheless, a good definition of life should precisely highlight and contribute to explain the link between the two sides³.

In order to do so, it is absolutely necessary to tackle the problem with a methodology that combines a ‘bottom-up’ approach (following the various stages in the process of emergence of living entities from complex physico-chemical phenomena) and a ‘top-down’ approach (specifying which is the final scenario of the journey: i.e., what may be already considered as ‘minimal life’). This reflects, again, the fact that the definition of life (i.e., the main ‘top-down’ constraint) and its origins (the ‘bottom-up’ sequence of events) are inseparable questions. Accordingly, our proposal is next articulated in two different sections: in the first (Section 4) we introduce and briefly explain our definition (which is meant to be well-rooted in present knowledge on living systems but, at the same time, is also meant to be a projection that captures their most universal properties); and in the second (Section 5) we offer a genealogical account of the cornerstones of this definition, pointing out the major transitions suggested as necessary and sufficient for the unfolding of any biological world. The basic mechanisms underlying each of those transitions will be indicated as we go along, so this second type of analysis will also serve to come out – at the end of that section – with an improved (more specific and operational) version of our definition.

4. Our Proposal: Autonomy and Open-ended Evolution as the Basic Ingredients of a New Universal Definition of Living Being

After reviewing two representative definitions in the field, it is our turn now to offer an alternative that, somehow, introduces advantages with respect to them. Even if these previous proposals roughly fulfill the requirements listed in section 2, we consider that the following definition does it in a more complete and satisfactory way (particularly concerning requirements ‘c’ – on conceptual elegance and explanatory power – and ‘e’ – on the balance between universality and specificity) and, therefore, means a significant improvement. We will start by defining life from an individual perspective (from the perspective of a single ‘living being’ or ‘organism’) and, later (especially in Section 6) we will discuss the implications of such a definition in a wider collective scale, introducing a conception of ‘life’ as a more encompassing, global phenomenon.

The new proposed definition: *‘a living being’ is any autonomous system with open-ended evolutionary capacities*, where

- (i) by *autonomous* we understand a far-from-equilibrium system that constitutes and maintains itself establishing an organizational identity of its own, a functionally integrated (homeostatic and active) unit based on a set of endergonic-exergonic couplings between internal self-constructing processes, as well as with other processes of interaction with its environment, and
- (ii) by *open-ended evolutionary capacity* we understand the potential of a system to reproduce its basic functional-constitutive dynamics, bringing about an un-

limited variety of equivalent systems, of ways of expressing that dynamics, which are not subject to any predetermined upper bound of organizational complexity (even if they are, indeed, to the energetic-material restrictions imposed by a finite environment and by the universal physico-chemical laws)⁴.

At first sight, this definition may seem too generic and not very operational, because it does not specify the type of components or molecular mechanisms that make an autonomous and evolutionary behavior of this kind possible (membrane, catalysts, energy currencies, informational records ...). To a great extent, that is the work we have ahead. In the next section we will analyze how this way of *being, doing* and *changing* in time can progressively be realized and integrated in a system (or group of systems). Through that analysis, we will be able to endow the concepts of 'autonomy' and 'open-ended evolution' with a more profound and accurate meaning, particularly looking into the connection between the two and discussing why one must come after the other⁵.

In any case, the present version already contains the fundamental (necessary and sufficient) theoretical ingredients to reconstruct the essential steps of the origins of life, from self-organizing physico-chemical phenomena until the constitution of systems with a level of complexity equivalent to that of the last universal common ancestor of all terrestrial forms of life. As we mentioned before, the autopoietic definition, by itself, would not lead us all the way up to that point (since an open-ended type of evolution requires the development of non-trivial mechanisms – e.g., those supporting hereditary reproduction), whereas the standard definition is not satisfactory because it does not provide the key conceptual tools to characterize the basic material organization required to get there ('self-maintenance' is, in that sense, much weaker and limited than 'autonomy')⁶.

5. Building a Genealogical Explanation of our Definition

The starting point of our account is a scenario in which self-organizing phenomena (complex physico-chemical phenomena like the so-called 'dissipative structures', Nicolis and Prigogine 1977) take place. By 'self-organization' we refer to a phenomenon occurring when a series of non-linear microscopic processes generate a global-macroscopic correlation (a new 'pattern of dynamic behavior') in far from equilibrium thermodynamic conditions that are maintained by the continuous action of a set of constraints, one of which – at least – is a result of the actual phenomenon.

It is quite clear that living beings cannot do without self-organizing processes (which are ubiquitous at all levels of biological organization), but the constraints that allow a living organism to stay in a far from equilibrium situation are much more elaborate. In fact, the organism *itself* elaborates the different constraints that control the flow of matter and energy through the system, unlike what happens

in the typical examples of dissipative structures (Bénard convection patterns, B-Z chemical reaction oscillations), where the flow that keeps the phenomenon running is externally harnessed. Thus, we can safely say that life is self-organization (we disagree with those who think that the ‘self’ should be dropped; see, for instance, Margulis and Sagan, 2002) but it is, indeed, much more than self-organization.

So, what is the main step after self-organization? Self-maintenance. Someone could say that in pure self-organizing phenomena it is already possible to start speaking about ‘self-maintenance’ (in so far as the generation of the macroscopic pattern contributes to its own maintenance by means of its continuous constraining action on the microscopic dynamics). However, in that case, we would be using the term in its weakest sense. A more significant *self*-maintenance cannot take place until a system starts producing some of the constraints that are crucial to control the matter-energy flow through it and, in this way, it begins to develop the capacity to maintain its organization in the face of external perturbations (i.e., a primitive kind of ‘organizational homeostasis’). A fact to be highlighted here is that this is only possible if the system is chemical, because the variety of constraining mechanisms required to achieve that capacity simply is not at reach for bare physical systems. Dynamics (understood in its classical sense: as the change in time of the position of a body or many-body system) is insufficient for such a task. Besides, as will be more obvious below, it is just through chemistry that a process of open-ended growth in complexity may take place⁷.

Thus, the next scenario is that of self-maintaining chemical networks. Systems made of relatively simple molecular components in which autocatalytic cycles are formed and have the potential to vary and grow in complexity (Kauffman, 1993; Wächtershäuser, 1988; de Duve, 1991) could be good candidates for this stage, provided that they are located in a proper – physically realistic – context: in between a source and a sink of free energy – as Morowitz (1968, 1992) rightly emphasizes through his ‘cycling theorem’. Thermodynamics matters. Although it is not possible to explain the relevant transitions at this stage just by means of thermodynamic tools and theories, it is very important not to forget about the implications of the general thermodynamic framework, which are certainly far-reaching. In particular, in the case of self-maintaining chemical networks it is crucial to consider the energetic requirements (both internal and with regard to the environment) for the reactions involved to actually take place: as we will see next, only those networks that are capable of establishing a set of endergonic-exergonic couplings between their constitutive processes will be candidates for basic autonomy (Ruiz-Mirazo and Moreno, 1998, 2000; Kauffman, 2000).

5.1. THE EMERGENCE OF BASIC AUTONOMY

The concept of self-maintenance, being important, still does not capture the most characteristic properties of a living organism: in particular, its capacity to build and rebuild continuously all the components and constraints that are responsible for its

organization and behavior, together with the capacity to adaptively modify that internal organization (plus the actual relation with the environment) as a response to external changes. Biological systems are component production systems (chemical networks) that manage their material and energetic resources in such a way that they continuously accomplish a global self-construction dynamics in a plastic way. But all this would not be possible if endergonic processes (non-spontaneous processes like the synthesis of polymers, or the transport of a substance against its electrochemical potential gradient) could not profit from the exergonic drive of other spontaneous processes. Thus, the actual generation – and regeneration – of the components that act as functional constraints in the system is based on the establishment of endergonic-exergonic couplings (and hence the deep entanglement between anabolism and catabolism in any metabolism).

We suggest that the transition from a bare self-maintaining chemical system to a full-fledged self-producing system with minimal adaptive capacities requires that the network of reactions gets encapsulated by a semi-permeable boundary of its own making and, at the same time, solves three fundamental problems: (i) an eventual osmotic crisis, (ii) the spatiotemporal coordination of all processes – and their couplings –, and (iii) the achievement of an efficient energy transfer through the components and processes of the system. The minimal set of components necessary to overcome these problems are (Ruiz-Mirazo, 2001): a *membrane* (an enclosure through which the system controls concentrations, establishes a clear cut distinction with the environment, and channels the interaction with it), a group of *energy currencies* (at least one soluble in water – like PP_i or ATP – and one directly related to transport processes – like the electrochemical potential gradient of protons or sodium ions across a membrane), and a set of *catalysts* (polymers – or initially just shorter oligomers like de Duve's (1991) 'multimers' – responsible for modulating the rates at which reactions take place, for setting up regulation/homeostatic mechanisms, and for carrying out mediated transport processes).

It is only when these three types of components get together in a single system and interrelate adequately that a complete enough set of endergonic-exergonic couplings may be accomplished and, together with it, a minimal autonomous system. It is important to notice that this dynamics involves both a new internal way of operating (functional relations among the components of the system)⁸ and a new way of interacting with the environment (agential behavior)⁹. Thus, this type of system can, in a basic and primitive sense, create its own 'umwelt'.

Some authors, like Luisi (1998), may argue that this kind of autonomous systems (cellular proto-metabolisms) already contain all that is required for life. Actually, it is difficult to say what we would do if we ever found such systems (in a test tube, on Mars or somewhere else): would we not call them 'living', despite the fact that they do not show a complex metabolic behavior based on macromolecules like genes or proteins? With regard to this point, we consider that basic autonomy is a fundamental property of life, but it is not the only one. In fact, the long-term maintenance of basic autonomous systems will not be guaranteed until they achieve

more metabolic versatility and robustness, mechanisms for reliable reproduction and heredity, together with an open-ended capacity to adapt and grow in complexity. In other words, the long-term stability of a biological world crucially depends on how living beings change in time and evolve through generations¹⁰. This is so much so that even what a living organism comes to be and do during its individual lifetime is profoundly influenced by the global evolutionary process in which it is inserted, as we will underline below.

In any case, we must point out that basic autonomy, although insufficient, is necessary for open-ended evolution. Without a component production machinery that solves the basic physical-material-energetic problems involved in the constitution of any (proto-)metabolic system it is impossible to reach the level of molecular and organizational complexity required to start a process of evolution of those characteristics. In addition, the constitution of basic autonomous systems also provides the necessary potential for functional diversification (i.e., a wide enough 'phenotypic space'), which is capital for the development of mechanisms of genetic information (Wicken 1987), and which is missing in models of molecular evolution that deal with populations of pure replicators (like typical 'RNA-world' models).

5.2. HEREDITARY AUTONOMOUS SYSTEMS: THE FIRST STEPS TOWARDS OPEN-ENDED EVOLUTION

Basic autonomous systems, as they have been described above, possess a fair degree of organizational homeostasis and adaptability, are capable of reproduction (even if it is just a 'statistical' kind of reproduction, by simple growth and division), and show a remarkable potential for diversification (because, in principle, they would be able to create ever-new components). However, they have no way of fixing or recording the novelties that appear in them (e.g., possible innovations in their metabolic dynamics) and, as their level of molecular and organizational complexity rises, their brittleness also increases.

According to some recent models that deal with relatively simple catalytic networks and study their growth and reproduction potential (see: Segré *et al.*, 2000), at this stage there could already be some transfer of 'information' from generation to generation (that is the idea behind what they call a 'compositional genome'). Of course, the actual autocatalytic dynamics that is responsible for growth (and subsequent division-multiplication) implies that some of the components and features of the 'mother system' get produced once and again, and possibly transmitted to the offspring. However, that continuous production of components and their transmission to the 'daughter cells' is not reliable. Sometimes everything will go fine and reproduction will be successful (i.e., mother and daughter cells will be equivalent, or have equivalent levels of complexity); but some others (a fraction of the whole that cannot be disregarded, especially if we deal with big populations and a big number of generations) the process will lead to unviable systems. And this is

a critical question when the complexity of the components and the difficulties to produce them become higher and higher.

Therefore, there is a major complexity bottleneck that can only be overcome if new, more sophisticated mechanisms of autocatalysis are developed by autonomous systems. These mechanisms must guarantee that the complex functional components of the system are well preserved in the on-going metabolic dynamics, as well as reliably transmitted to the next generations through subsequent processes of reproduction. The solution is to generate functional components that perform ‘*template activity*’: i.e., material structures that can be faithfully *replicated*¹¹, regardless of their complexity (i.e., at this stage: the length of the polymer chain) and of the frequency at which the process is carried out. In this way, components become material ‘records’ and it is possible to start speaking about ‘heredity’ in a more rigorous way (Pattee, 1967, 1982, 1997).

As a consequence, a new scenario emerges with systems whose metabolism is more complex and robust than before, whose reproduction is more reliable, and whose evolutionary dynamics, apart from being more selective (stronger pressure among synchronous competitors for limited resources), already introduces the possibility of defining – and tracing – ‘lineages’ (i.e., asynchronous relations between systems with non-coherent lifetimes). This stage of ‘hereditary autonomous systems’ (systems with an operational organization based on a single type of functional polymer¹²) is analogous to an ‘RNA-world’ but, of course, provided that such a ‘world’ is embedded in a cellular metabolism, and not simply set up experimentally or theoretically in the context of a flow reactor, as it is commonly done (for a review see Eigen, 1992).

5.3. THE GENOTYPE-PHENOTYPE DECOUPLING AND THE ORIGINS OF OPEN-ENDED EVOLUTION

In any case, hereditary autonomous systems face a serious limitation when it comes to initiating a truly open-ended evolutionary process: the development of mechanisms that allow for a more and more reliable conservation and transmission of components from generation to generation turns out to be structurally incompatible with the development of the functional efficiency and metabolic versatility of those components. This happens because – as we just mentioned above – these are systems based on a single type of complex functional component, which articulates both the *realization* and the *evolution* of their metabolism, and which needs to accommodate simultaneously two very different properties: template activity and catalytic specificity/efficiency (Moreno and Fernández, 1990; Benner, 1999).

The problem has a rather simple chemical interpretation. Template activity requires a stable, uniform morphology, suitable to be lineally copied (i.e., a monotonous spatial arrangement that favors low reactivity and it is not altered by sequence changes); whereas catalytic diversity requires precisely the opposite: a very

wide range of three-dimensional shapes (configuration of catalytic sites), which are highly sensitive to variations in the sequence.

This ‘trade-off’ problem marks out a second major bottleneck in the way to establish the conditions that permit an indefinite growth of complexity. Such a bottleneck can only be overcome by those autonomous systems which start producing *two* different types of polymers: some ‘new records’ specialized in the reliable recording, storage and replication of certain linear structures – or polymer ‘sequences’ – that become crucial to ensure their correct functioning (and that of similar systems in future generations); and some others, the ‘new catalysts’, specialized in carrying out with increasing efficiency the strictly metabolic tasks required for their continuous realization as individual autonomous systems.

The new records would be, somehow, *decoupled* from metabolism, in quasi-inert states, and would have hardly any causal effects on the system’s functioning if they did not act as a fundamental *reference* for the new catalysts (Pattee, 1982; Moreno and Ruiz-Mirazo, 2002). More accurately, the former must *instruct* the synthesis of the latter: i.e., records must become responsible for the correct specification of catalysts. And these, in turn, must control and catalytically regulate all the processes in which records are involved (replication, translation, reparation, etc.), although without taking part directly in the generation and/or alteration of those records (since this should be determined by the global evolutionary dynamics of the population of autonomous agents).

The key lies, therefore, on the establishment of a certain circularity – a causal connection or correlation – between those two types of components (and the processes associated to each of them). But, given the structural incompatibility between the capacity for an accurate template replication and the development of more proficient catalysts, it is virtually impossible that the connection between the two could be built on their intrinsic properties as polymers. On the contrary, the conditions from which those components stem (i.e., the previous single-polymer world) relations establish through a series of intermediate-components, i.e.: in a non-linear and clearly *indirect* way (even if the potential to have causal effects on one another is not lost). Such an indirect, mediated connection is, in fact, a requisite to achieve an effective decoupling between those two operational modes or levels in the system. And this is precisely what is required for the emerging type of metabolic organization (which may be called ‘instructed metabolism’) to combine coherent and fruitfully the two new dimensions of its activity: the (individual) ontogenic-functional one and the (collective) phylogenic-evolutionary one.

The most natural way of understanding this new ordering in the system, this new (partially decoupled but strongly complementary – i.e., *indirectly coupled* –) organizational structure is through the idea of ‘information’¹³. Other authors have defended a similar position, holding that what distinguishes life is precisely the fact that there is a code-type relation between template and functional components (Pattee, 1977, 1982; Hoffmeyer and Emmeche, 1991; Umerez, 1995).

It is only once hereditary autonomous systems start producing ‘informational’ components and mechanisms (i.e., once a *translation code*¹⁴ appears between two very different types of functional components in the system) that the ‘genotype-phenotype’ distinction becomes really significant (even if some authors – for example, Eigen, 1992 – try to anticipate it) and we can start speaking properly about open-ended evolution. This is so because only through the generation of components which can be replicated with very high reliability, which have energetically degenerate states and, most importantly, which are interpreted as instructions by the system (i.e., ‘informational records’ – or ‘unlimited hereditary replicators’ in Maynard Smith and Szathmáry’s (1995) terminology) can autonomous systems begin to construct *indefinitely* new forms of molecular and organizational complexity.

5.4. COROLLARY: THE DEFINITION REVISITED

In sum, in order to arrive at a fully-fledged biological scenario, it is necessary to go all the way up until some genetically instructed collection of metabolisms appear. If we stopped before (e.g., at the stage of basic or hereditary autonomous systems) we would not have all the key ingredients that support the organization of a minimal living being, and we would have to face the problem of needing two different – and incompatible – conceptions of minimal life. For the selection of those key ingredients must be done taking into account two rules: the ‘continuity thesis’ between the inert and the living worlds, and the fact that they must be mechanisms or features of the organization that are so fundamental that cannot be erased from then onwards, throughout the subsequent evolutionary history of that type of organization (as it has happened throughout the evolutionary history of life on Earth).

Accordingly, we can make our original definition of living being as ‘an autonomous system with open-ended evolutionary capacities’ more explicit and operational, since it is possible to state now that, in order to accomplish those two properties, any system must have:

a semi-permeable active boundary (i.e., a membrane), an energy transduction/conversion apparatus (a set of energy currencies) and, at least, two types of interdependent macromolecular components: some carrying out and coordinating directly self-construction processes (catalysts) and some others storing and transmitting information which is relevant to carry out efficiently those processes in the course of subsequent generations (records).

6. Theoretical Implications: Life is A Historical-collective Phenomenon

Probably the most important theoretical conclusion that can be drawn from the last section is that a living being cannot exist but in the context of a global network of similar systems¹⁵. This is reflected in the fact that informational components (which specify the metabolic components and organization of single biological entities), in order to be functional, have to be shaped through a process that involves a great amount of individual systems and also very many consecutive generations (or reproductive steps). Such a collective process, which has both a synchronic-ecological side and a diachronic-evolutionary one, is actually crucial for the sustainability of the living phenomenon as a whole.

On the one hand, since the material resources of any real physical environment – like a planet – are limited, life must learn how to make the best use of what is available and also of what it continuously produces. In the individual sphere, this means that living agents have to compete for organic compounds or its precursors; so, under selective pressure, depredation strategies are bound to appear and, thereafter, full ‘food webs’, too (initially, of course, just between autotrophic and heterotrophic unicellular systems). In the collective sphere, the ecosystem eventually has to deal with the problem of recycling bioelements at a global scale. Otherwise, there would be a major crisis that would put at risk the persistence of that whole biological world. The solution is to couple with geo-physical and geo-chemical processes that take place on the planet and establish global bio-geo-chemical cycles. As life cybernetically controls the environmental conditions of the inhabited planet, such as the composition of the atmosphere (Lovelock and Margulis, 1974a, b), this will allow for its long-term maintenance.

On the other hand, an open-ended process of evolution (which does not mean any evolutionary arrow towards higher and higher complexity, but just the possibility that there is a steady increase in it) is precisely what allows the continuous renewal of particular individual organisms and particular types of organism for the sake of keeping the overall process running. This open-ended historical process takes place at a much longer time-scale than the typical lifetimes of individual systems and – as we said in the previous section – strictly speaking, it only begins when those systems are endowed with a genetic machinery that instructs their metabolism and is transmitted reliably through generations. Nevertheless, that evolutionary dynamics of the ecosystems does not appear ‘off the cuff’: it must be progressively articulated during stages previous to the origins of genetics (i.e., previous to the origins of life).

In other words, genetic information is the central concept to understand the intricate connection between the functional dynamics of individual organisms and the evolutionary dynamics of the global biological network. However, since that connection cannot be suddenly established, we have to focus the analysis on the *origins* of genetic information (i.e., on which are the most plausible conditions for the emergence of genetic information). And that leads us to study the nature of

chronic process), generates new records – genetic information – which, once embedded in the individual organizations, allow for new, open-ended, functional interactions.

There is such a deep entanglement between these two levels that both the collective and the individual organization of life are cause and consequence of each other. Nevertheless, it is important to underline that there is also a basic asymmetry between the individual (metabolic) network and the collective (ecological) one: both are self-maintaining and self-producing organizations, *but* only individual living beings (organisms) are autonomous *agents* with a self-produced, active physical border, plus a high degree of *functional integration* among components, plus a machinery for *hereditary reproduction*.

7. Practical Implications

As a final conclusion to this article, we would like to highlight briefly the main practical implications of our approach and of the general conception that has been offered here for future research. In this sense, four different avenues of research are considered:

- (a) *Origins of Life on Earth*. Present research is strongly biased by molecular biology and genomics, and so it focuses on the main structural features, relational properties and possible abiotic synthetic roots of biopolymers (proteins, RNA, DNA) and its precursors (for review see Lazcano, 2001). Our proposal, based on the conviction that ‘energetics’ is to be worked out before ‘genetics’ can take over (Morowitz, 1981) (and, thus, that a proto-metabolic organization is required before complex biopolymers can arise; Morowitz *et al.*, 1988; de Duve, 1991), favors an alternative paradigm: the search into different self-constructing cellular systems, that may later lead to the production of functional macromolecules (always to be regarded in the context of the whole metabolic organization). Our definition allows, in any case, to signalize the general sequence of transitions from the non-living to the living (see Figure 1).
- (b) *Astrobiology*. The definition suggested in this article (particularly the extended version) provides the necessary conceptual framework to develop biomarkers, as well as criteria to tell the living from the inert material organizations. (i) At an individual level, a living system must have a semi-permeable active boundary, a set of energy currencies and, at least, two types of interdependent macromolecular components (analogous to proteins and nucleic acids); therefore, a virus would not be living, but a sterile organism would. (ii) At the global level, life would only exist as a long-term collective-historical phenomenon on a planet, which would be made long-term habitable as a consequence of

the spatial unfolding of life and the operation of its homeorretic mechanisms (Sagan and Margulis, 1984; Lovelock, 1988). Accordingly, there should be external signs (like the strongly non-equilibrium composition of the atmosphere) that would tell us if there is life or not on the planet (Hitchcock and Lovelock, 1967).

- (c) *Artificial Life*. From the three main sub-fields that Artificial Life has developed into (i.e.: purely computational models, robotics and *in vitro* experiments), at the moment the most congruent with our perspective is the *in vitro* approach (Deamer, 1998; Szostak *et al.*, 2001). Computational models tend to disregard the physical-material-energetic aspects of biological organization (which we consider crucial), while robotics – even the so called ‘autonomous robotics’ – pays very little attention to the problem of self-construction (which is, indeed, rather difficult – if not impossible – to deal with outside chemistry). Thus, we consider that the most promising avenue of research in this area is that of *in vitro* experiments, provided that they are primarily focused on the chemical implementation of relatively simple cellular proto-metabolisms (i.e., basic autonomous systems), and specially if they do so without borrowing components or mechanisms from existing biological systems (Bro, 1997). This challenge could nowadays be at reach for us (even if bridging the gap from such basic autonomous systems to life – under realistic experimental conditions – will be quite a different issue).
- (d) *Biological Theory*. We consider that a profound conceptual debate and clarification is required for the development of a general theory of biology. Attempts to reach consensus on a possible universal definition of life could be very helpful for that task. Our proposal is meant to be a contribution in this direction, putting forward autonomy and open-ended evolution as the key ideas to work on. No general theory of biological organization will be ready until we discover the mechanisms and principles underlying those basic properties and, specially, the way they interweave to produce the living phenomenology.

Notes

¹ Contrast with the ‘seven pillars’ of de Duve (1991): Manufacture of the own components, extract energy from the environment, catalysis, inform cell processes, insulation, regulation and multiplication.

² Other theoretical models that have addressed the problem of which could be the minimal living organism in general-abstract terms (like Rosen’s approach or Ganti’s ‘chemoton theory’: see, for instance, Rosen, 1991; Ganti, 1987) are subject to a similar type of criticism. Any metabolic or proto-metabolic dynamics really takes place in far from equilibrium conditions (and, in order to keep those conditions, all possible minimal organisms must form – and be understood – as thermodynamically open systems). Thus, it is crucial to introduce the energetic-thermodynamic dimension of the problem from the very beginning.

³ This can be achieved even if the definition is formulated from the point of view of individual organisms, i.e., in terms of what a 'living being' is and how it organizes and behaves (like authors of the autopoietic school do). The crucial point (which those authors, however, do not consider relevant) is to include in the definition the main feature(s) that will reveal, at the individual-metabolic level of analysis, the existence – and the important implications – of being inserted in a collective-ecological-evolutionary dynamics. Our proposal, in fact, will be elaborated on these lines.

⁴ According to this criterion, what is really crucial for a system to be considered biological is to have the *potential* to be part of a process of open-ended evolution. In fact, there are living organisms that in natural conditions are not able to reproduce (i.e., sterile organisms, like mules, working bees or plants without seeds) and, thus, cannot continue on the evolutionary process that actually brought them about. Nevertheless, the analysis of the molecular components of any of those organisms would reveal that they have all the mechanisms required for that type of evolution (see below, particularly the end of Section 5). Ontogenic constraints prevent their natural reproduction and their subsequent participation in such a process (even if we could easily think of various ways to overcome those difficulties – through artificial cloning or vegetative reproduction, for instance). Therefore, through the use of the term 'capacity' in the second part of our definition we introduce a certain tension between the 'actual' and the 'in-principle-accessible' properties of living beings, which resembles Ganti's (1987) distinction between *real (absolute)* and *potential* criteria for life, and which is necessary – as he showed – to avoid contradictory results.

⁵ The reader might associate these two fundamental concepts with the names of Varela (1979) and Von Neumann (1966), respectively (see also McMullin, 2000). However, our formulation is radically different because it is intended to capture the physical-material-energetic requirements behind those concepts (whereas Varela's or Von Neumann's approaches are far more abstract).

⁶ Despite the apparent formal similarity with the 'standard definition', our proposal is more restrictive, more demanding and, at the same time, it has a higher explanatory potential. As will be shown with more detail in the next section, autonomy requires much more than the temporary self-maintenance of a chemical network.

⁷ This issue (i.e., chemistry's relevance for the development of any biological world) would deserve a much longer discussion but, since it is not the central focus of the present article, we just offer a very condensed line of argument. We can add in this brief note that the processes that allow the true *self-maintenance* of a system in far-from-equilibrium conditions must be processes of transformation of matter (i.e., chemical processes), because the components of such a system (including the components that would have to do more directly with the inter-phase or boundary conditions of the system – i.e., with the potential control of the flow of matter-energy through the system) are not stable out of that continuous 'turn-over' *chemical* dynamics.

⁸ Functional relations appear when all the components of a system contribute to or participate in the maintenance of such a system.

⁹ This 'basic agency' has to do with the capacity of the system to control actively some of its boundary conditions – like concentration gradients.

¹⁰ Perhaps the exploration of Mars could show traces of proto-metabolisms, of basic autonomous systems, which did not (or have not) overcome the critical transitions to start an open-ended evolutionary dynamics and, therefore, did not (or have not) achieve(d) their long-term maintenance on the planet.

¹¹ Following Dyson (1985), Fleischaker (1994) and Luisi (1994), we distinguish the terms ‘replication’ and ‘reproduction’: Replication is a reliable copying process that takes place at a molecular level, whereas reproduction involves the spatial multiplication of a whole organization.

¹² A type of polymer that can already evolve through mutational variations in its space of possible sequences (Benner, 1999).

¹³ By this term we mean a kind of causal connection in a system by which some (quasi) inert material patterns constrain, through a certain mechanism of ‘translation-interpretation’, the metabolic dynamics of the system. In turn, it is only through the participation of the metabolic machinery of the very system that the informational patterns become instructions (see Moreno *et al.*, 1994; Moreno and Ruiz-Mirazo, 2002).

¹⁴ Like the ‘genetic code’ between DNA and proteins present in all living beings.

¹⁵ This is fully coherent with recent work that criticizes the traditional idea of ‘common ancestor’ (Doolittle, 1999) and suggests that, instead of the common image of *a cell* from which all other living entities on Earth derived (through ‘vertical evolution’), we should think in terms of a ‘common ancestral population’ of very similar metabolic systems in constant evolution (both ‘vertical’ and ‘horizontal’).

¹⁶ Process of constitution of an ecosystem (Haynes, 1990).

Acknowledgements

This article was presented at the Oaxaca ISSOL’02 Meeting, 30 June–5 July, 2002. The authors are grateful to Dr. Jon Umerez for many helpful discussions. K. R.-M. was holding a fellowship from the Instituto Nacional de Técnica Aeroespacial (INTA) of Spain at the Centre of Astrobiology (Madrid) during the preparation and writing of this article. The work was also supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT) of Spain, Grant BMC2000-0764, and by the University of the Basque Country (UPV/EHU) Grant 9/upv 00003.230-13707/2001.

References

- Benner, S. A.: 1999, How Small can a Microorganism be? in *Size Limits of Very Small Microorganisms (Proceedings of a Workshop)*, National Academy Press, Washington, D.C., pp. 126–135.
- Bro, P.: 1997, Chemical Reaction Automata, *Complexity* **2**, 38–44.
- Brooks, D. R. and Wiley, E. O.: 1988, *Evolution as Entropy: Toward a Unified Theory of Biology*, University of Chicago Press, Chicago (reviewed edition of 1986).
- Cleland, C. E. and Chyba C. F.: 2002, Defining ‘Life’, *Origins Life Evol. Biosph.* **32**, 387–393.
- de Duve, C.: 1991, *Blueprint for a Cell: The Nature and Origin of Life*, Neil Patterson, Burlington, p. 4.
- Deamer, D. W.: 1998, ‘Membrane Compartments in Prebiotic Evolution’, in A. Brack (ed.), *The Molecular Origins of Life. Assembling the Pieces of the Puzzle*, Cambridge University Press, Cambridge, pp. 189–205.

- Doolittle, W. F.: 1999, Phylogenetic Classification and the Universal Tree, *Science* **284**, 2124–2128.
- Dyson, F. J.: 1985, *Origins of Life*, Cambridge University Press, Cambridge.
- Eigen, M.: 1992, *Steps Towards Life: A Perspective on Evolution*, Oxford University Press, New York.
- Eigen, M. and Schuster, P.: 1979, *The Hypercycle: A Principle of Natural Self-organization*, Springer, New York.
- Emmeche, C.: 1998, Defining Life as a Semiotic Phenomenon, *Cybernet. Human Knowing* **5**, 3–17.
- Farmer, J. D. and Belin, A. d'A.: 1992, 'Artificial Life: The Coming Evolution', in C. G. Langton, C. Taylor, J. D. Farmer and S. Rasmussen (eds), *Artificial Life II*, Santa Fe Institute Studies in the Sciences of Complexity Proceedings, Vol. X, Addison-Wesley, Redwood City, CA, pp. 815–838.
- Fernández, J., Moreno, A. and Etxeberria, A.: 1991, Life as Emergence: The roots of a New Paradigm in Theoretical Biology, *World Futures (The Journal of General Evolution)* **32**, 133–149.
- Fleischaker, G. R.: 1988, Autopoiesis: The Status of its System Logic, *BioSystems* **22**, 37–49.
- Fleischaker, G. R.: 1994, 'A Few Precautionary Words Concerning Terminology', in G. R. Fleischaker, S. Colonia and P. L. Luisi (eds), *Self-production of Supramolecular Structures*, Kluwer Academic Publishers, Dordrecht, pp. 33–41.
- Ganti, T.: 1987, *The Principle of Life*, OMIKK, Budapest.
- Haynes, R. H.: 1990, 'Ecce Eco-poiesis: Playing God on Mars', in D. MacNiven (ed.), *Moral Expertise: Studies in Practical and Professional Ethics*, Routledge, London, pp. 161–183.
- Hitchcock, D. R. and Lovelock, J. E.: 1967, Life detection by atmospheric analysis, *Icarus* **7**, 149–159.
- Hoffmeyer, J. and Emmeche C.: 1991, 'Code-duality and the Semiotics of Nature', in M. Anderson and F. Merrell (eds), *On Semiotic Modeling*, Mouton de Gruyter, New York, pp. 117–166.
- Joyce, G. F.: 1994, 'Foreword', in D. W. Deamer and G. R. Fleischaker (eds), *Origins of Life: The Central Concepts*, Jones and Bartlett, Boston, pp. xi–xii.
- Kauffman, S.: 1993, *The Origins of Order: Self-organization and Selection in Evolution*, Oxford University Press, Oxford.
- Kauffman, S.: 2000, *Investigations*, Oxford University Press, Oxford, Ch. 3.
- Koshland Jr., D. E.: 2002, The Seven Pillars of Life, *Science* **295**, 2215–2216.
- Lazcano, A.: 2001, 'Origin of Life', in D. E. G. Briggs and P. R. Crowther (eds), *Paleobiology II*, Blackwell Science, Oxford, pp. 3–8.
- Lewontin, R. C.: 1970, The Units of Selection, *Annu. Rev. Ecol. System.* **1**, 1–18.
- Lovelock, J. E.: 1988, *The Ages of Gaia*, Norton, New York.
- Lovelock, J. E. and Margulis, L.: 1974a, Atmospheric Homeostasis by and for the Biosphere: The Gaia Hypothesis, *Tellus* **26**, 2–10.
- Lovelock, J. E. and Margulis, L.: 1974b, Homeostatic Tendencies of the Earth's Atmosphere, *Origins of Life* **1**, 12–22.
- Luisi, P. L.: 1998, About Various Definitions of Life, *Orig. Life Evol. Biosph.* **28**, 613–622.
- Luisi, P. L.: 1994, 'The Chemical Implementation of Autopoiesis', in G. R. Fleischaker, S. Colonna and P. L. Luisi (eds), *Self-production of Supramolecular Structures*, Kluwer Academic Publishers, Dordrecht, pp. 179–197.
- Margulis, L.: 1990, 'Big Trouble in Biology: Physiological Autopoiesis Versus Mechanistic Neo-Darwinism', in J. Brockman (ed.), *Doing Science: The Reality Club 2*, Prentice Hall, NY, pp. 211–235.
- Margulis, L. and Sagan, D.: 2002, *Acquiring Genomes. A Theory of the Origin of Species*, Basic Books, New York, Ch. 2.
- Maturana, H. and Varela, F. J.: 1973, *De Máquinas y seres Vivos – Una Teoría Sobre la Organización biológica*, Editorial Universitaria S.A., Santiago de Chile.
- Maynard Smith, J.: 1986, *The Problems of Biology*, Oxford University Press, Oxford.
- Maynard Smith, J. and Szathmáry, E.: 1995, *The Major Transitions in Evolution*, Freeman and Co., Oxford.

- Mayr, E.: 1982, *The Growth of Biological Thought*, Harvard University Press, Cambridge, MA.
- McMullin, B.: 2000, John von Neumann and the Evolutionary Growth of Complexity: Looking Backward, Looking Forward ..., *Artificial Life* **6**, 347–361.
- Moreno, A. and Fernández, J.: 1990, Structural Limits for Evolutive Capacities in Molecular Complex Systems, *Biology Forum* **83**, 335–347.
- Moreno, A. and Ruiz-Mirazo, K.: 1999, Metabolism and the Problem of its Universalization, *BioSystems* **49**, 45–61.
- Moreno, A. and Ruiz-Mirazo, K.: 2002, Key Issues Regarding the Origin, Nature and Evolution of Complexity in Nature: Information as a Central Concept to Understand Biological Organization, *Emergence* **4**, 63–76.
- Moreno, A., Umerez, J. and Fernandez, J.: 1994, Definition of Life and Research Program in Artificial Life, *Ludus Vitalis* **2**, 15–33.
- Morowitz, H. J.: 1968, *Energy Flow in Biology*, Academic Press, New York.
- Morowitz, H. J.: 1981, Phase Separation, Charge Separation and Biogenesis, *BioSystems* **14**, 41–47.
- Morowitz, H. J.: 1992, *Beginnings of Cellular Life*, Yale University Press, New Haven.
- Morowitz, H. J., Heinz, B. and Deamer, D. W.: 1988, The Chemical Logic of a Minimum Protocell, *Origins Life Evol. Biosph.* **18**, 281–287.
- Nicolis, G. and Prigogine, Y.: 1977, *Self-organization in Non-equilibrium Systems*, Wiley, New York.
- Oparin, A. I.: 1961, *Life. Its Nature, Origin, and Development*, Academic Press, New York, Ch. 1.
- Pattee, H. H.: 1967, Quantum Mechanics, Heredity and the Origin of Life, *J. Theor. Biol.* **17**, 410–420.
- Pattee, H. H.: 1977, Dynamic and Linguistic Modes of Complex Systems, *Int. J. Gen. Syst.* **3**, 259–266.
- Pattee, H. H.: 1982, Cell Psychology: An Evolutionary Approach to the Symbol-matter Problem, *Cognit. Brain Theor.* **4**, 325–341.
- Pattee, H. H.: 1997, ‘The Physics of Symbols and Evolution of Semiotic Controls’, in M. Coombs (ed.), *Workshop on Control Mechanisms for Complex Systems: Issues of Measurement and Semiotic Analysis*, New Mexico State University.
- Pályi, G., Zucchi, C. and Caglioti, L.: 2002, *Fundamentals of Life*, Elsevier, Paris.
- Rosen, R.: 1991, *Life Itself: A Comprehensive Inquiry into the Nature, Origin and Fabrication of Life*, Columbia University Press, New York.
- Ruiz-Mirazo, K.: 2001, ‘Condiciones Físicas para la Aparición de Sistemas Autónomos con Capacidades Evolutivas Abiertas’, *Ph.D. Dissertation*, University of the Basque Country, San Sebastián.
- Ruiz-Mirazo, K. and Moreno, A.: 1998, ‘Autonomy and Emergence: how Systems become Agents through the Generation of Functional Constraints’, in G. L. Farre and T. Oksala (eds), *Emergence, Complexity, Hierarchy, Organization* (Selected and edited papers from the ECHO III Conference), *Acta Polytech. Scand.* **Ma91**, The Finnish Academy of Technology, Espoo-Helsinki, pp. 273–282.
- Ruiz-Mirazo, K. and Moreno, A.: 2000, Searching for the Roots of Autonomy: The Natural and Artificial Paradigms Revisited, *J. Integr. Study Artific. Intel. Cogn. Sci. Appl. Epistemol.* Special Issue on Autonomy **17**, 209–228.
- Sagan, C.: 1970, ‘Life’, in *The Encyclopaedia Britannica*, William Benton, London.
- Sagan, D. and Margulis, L.: 1984, ‘Gaia and Philosophy’, in L. S. Rouner (ed.), *On Nature*, University of Notre Dame Press, Notre Dame, IN, pp. 60–75.
- Segré, D., Ben-Eli, D. and Lancet, D.: 2000, Compositional Genomes: Prebiotic Information Transfer in Mutually Catalytic Non-covalent Assemblies, *Proc. Nat. Acad. Sci. U.S.A.* **97**, 4112–4117.
- Shapiro, R.: 2002, ‘Monomer World’, in *Abstracts of the 13th International Conference on the Origin of Life, 10th ISSOL Meeting*, Oaxaca, Mexico, p. 60.

- Shapiro, R. and Feinberg, G.: 1990, 'Possible Forms of Life in Environments very Different from the Earth', in J. Leslie (ed.), *Physical Cosmology and Philosophy*, MacMillan, New York, pp. 248–255.
- Szostak, J. W., Bartel, P. and Luisi, P. L.: 2001, Synthesizing Life, *Nature* **409**, 387–390.
- Umerez, J.: 1995, 'Semantic Closure: A Guiding Notion to Ground Artificial Life', in F. Morán, A. Moreno, J. J. Merelo and P. Chacón (eds), *Advances in Artificial Life*, Springer-Verlag, Heidelberg, pp. 77–94.
- Varela, F. J.: 1979, *Principles of Biological Autonomy*, Elsevier, New York.
- Varela, F. J.: 1994, 'On Defining Life', in G. R. Fleischaker, S. Colonna and P. L. Luisi (eds), *Self-production of Supramolecular Structures*, Kluwer Academic Publishers, Dordrecht, pp. 23–31.
- Von Neumann, J.: 1966, *Theory of Self-reproducing Automata*, A. W. Burks (ed.), University of Illinois, Urbana.
- Wächtershäuser, W.: 1988, Before Enzymes and Templates: Theory of Surface Metabolism, *Microbiological Reviews* **52**, 452–484.
- Wicken, J. S.: 1987, *Evolution, Thermodynamics and Information. Extending the Darwinian Program*, Oxford University Press, Oxford.