HAL
open science

The definition of life in the context of its origin

Y. N. Zhuravlev, V. A. Avetisov

To cite this version:

Y. N. Zhuravlev, V. A. Avetisov. The definition of life in the context of its origin. Biogeosciences, 2006, 3 (3), pp.281-291. hal-00297562

HAL Id: hal-00297562

https://hal.science/hal-00297562v1

Submitted on 18 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Biogeosciences, 3, 281–291, 2006
www.biogeosciences.net/3/281/2006/
© Author(s) 2006. This work is licensed
under a Creative Commons License.

Biogeosciences

The definition of life in the context of its origin

Y. N. Zhuravlev[1] **and V. A. Avetisov**[2]

[1]Institute of Biology and Soil Science, Russian Academy of Sciences, Far Eastern Branch, 100-letia 159, 690022 Vladivostok, Russia
[2]The Semenov Institute of Chemical Physics, Russian Academy of Sciences, Kosygina 4, 119991 Moscow, Russia

Received: 5 October 2005 – Published in Biogeosciences Discuss.: 10 February 2006
Revised: 23 May 2006 – Accepted: 31 May 2006 – Published: 10 July 2006

Abstract. Current life is a complex, multi-level phenomenon that is so diverse in its manifestations that a short and exhaustive definition of life is hardly possible. The high complexity of life, as well as a poor understanding of what life is in essence, are the obstacles to the elaboration of such a definition. Important characteristics of life, such as whole system-, ecosystem-, and information-defined characteristics, have been included in the definition of life only recently. Ecosystem-defined characteristics have been absent in models of the pre-biotic state for a long time. However, without an ecosystem context, the concept of the emergence of life cannot be complete. Interconnections between living and non-living components of a primordial evolving system are decisive for the period of transition from chemical to biological evolution.

Information-defined characteristics of life are often reduced to storage and the expression of genetic information, yet the operation of such perfect processes in prebiotic and transitional systems is unlikely. Genetic information, as defined in terms of the Shannon theory of communication, represents only a certain "informational channel" specified with respect to the expression of the structural genes. However, recent findings concerning the molecular mechanisms of the differential regulation of gene activity, and in the genomics, postgenomics and proteomics control mechanisms, suppose a richer diversity of informational flows in the organism. Moreover, considering life in a more general context, other types of related, informational channels, in particular, regarding the differentiation of higher taxa, hiatus, and expansion processes, should be kept in mind.

In many publications devoted to the origin of life, the terms "living", "life", and "living organism" are freely interchanged which proves the vagueness of insights about the different levels of the living system.

This report considers some variants of the definition of life that have been recently suggested and are based on present-day knowledge of the structures and functions of life. The contradictory demands of a definition, which needs to be complete and short at the same time, are emphasized. A definition characterizing life as a state, a structure, and a process, is proposed.

Correspondence to: Y. N. Zhuravlev
(zhuravlev@ibss.dvo.ru)

1 Introduction

Our insight into the present-day reality is circumscribed and fragmentary. It is rather a complicated goal to define even a familiar event or phenomenon, and the phenomenon of life is especially pertinent in this respect. A large number of definitions of life have been proposed (about 80 short definitions were presented in the proceedings of the International Workshop on Life, Pályi et al., 2002), some of them being controversial and none was in common use. Nevertheless, such attempts are not meaningless, since in trying to define the object we obtain more and more comprehensive illumination of the problem and can continue such attempts as long as the phenomenon is revealed. However, in defining life, many authors find this point unreachable and even Bohr (1933) supposed that "the existence of life must be considered as an elementary fact (or axiom) that cannot be explained, but must be taken as a starting point in biology". Similar problems arise when we try to reconstruct the processes that resulted in the emergence of life. The reconstructions are more difficult since the events leading up to the origin of life are remote from us by a time interval of about 4 billions years and are not preserved in the early geological record.

It seems unlikely that by combining two abstruse problems (the definition of recent life and the reconstruction of its origin) we can cast additional light on either of them. However, these two terms (the definition of life and the origin of life) can serve as a mutual test for the validity of hypotheses in

Table 1. Different notions of life emergence and their scenarios.

Notion or belief	Scenario
- Life is eternal	- No
- Life is created by God	- Hexaemeron or similar
- Life emerged via evolution from preexisting more simple organic matter	- Reconstruction of a transition from chemical to biological evolution
- Life invaded from space (panspermia)	- Any of three above

both fields, and, on the other hand, different scenarios for the emergence of life can be proposed depending on the core attributes of life chosen for its definition. The latter becomes especially evident when different notions of the emergence of life are summarized in Table 1. The two first lines of the table present beliefs which need no scientific reconstructions, the fourth one drives us back to the three possibilities mentioned above, and only the third notion is based on the idea of scientific reconstruction.

The third scenario (Table 1) is presented in a general form. The "metabolism first" scenario can be subdivided into the "prebiotic soup" (Oparin, 1924; Haldane, 1929), the "surface metabolism" (Wächtershauser, 1988) hypotheses, and other variants of the chemoorganotrophic concept (for references, see Wächtershauser, 1988). Similarly, the "two polymers" scenario includes "the gene-first" (Gilbert, 1986; see also more recent references in Santos et al., 2003), as well as "the replication-first" scenario comprising the hypercycle and its stochastic versions (Eigen and Schuster, 1979; Eigen et al., 1981; Szathmáry and Demeter, 1987). Among these concepts, the RNA-world is the most accepted (Gilbert, 1986; Gesteland et al., 1999), although it is also criticized (Orgel, 2003). The Lipid-world (Segre et al., 2001) is a less popular scenario of pre-life, nevertheless it should also be mentioned here.

We have insufficient scientific evidence to prefer any of these scenarios. Moreover, it cannot be excluded that some of the scenarios were not implemented in concert. However, almost every author believes that his idea based upon personal knowledge, experience, and persuasions, is more fundamental than others. In addition to these personal preferences, some objective circumstances influence the content and sense of the definition. For example, an observer on Earth will discover such attributes of life as the diversity of species in nature and diversity of domesticated forms, competition and synergism of species; these observations impelled Darwin (1872) to put forward the ideas of speciation and evolution. On the contrary, the above listed important attributes are indiscernible for an observer positioned outside the Earth, and he will define life as a process of current expansion by energy and matter sources from the surroundings.

In this paper, we consider the life coming-to-be as a transition from chemical evolution to a biological one. Such inseparable attributes of life as the transformation of energy and matter were inherited by life from the preceding stages of the Earth's development, whereas complexity, active adaptation, and multi-level hierarchy were established directly during the transition period. In general, the transition period should have been rather short, but extremely rich in events, resulting in the life's intrinsic characteristics.

2 Historical variations of the definition of life

There is a set of claims which any definition has to satisfy, and this general problem has been discussed in detail elsewhere (Emmeche, 1998; Cleland and Chyba, 2002; Ruiz-Mirazo et al., 2004). According to these authors 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.).

One can notice that multi-sided requirements accumulated in these five items are very advisable, yet hardly achievable in the framework of one elegant concept. For example, Koshland's (2002) "seven pillars" proposal is rather elegant. Yet, when extrapolated to primordial life, the "pillars" suffers from differential rescaling that destroys the whole building. Indeed, it seems unlikely that properties of "compartmentalization, program and seclusion" are the features of non-differentiated primary life. So elegance is, most likely, a satellite of the incompleteness or uncertainty of definition, rather than of its deep explanatory power.

The majority of definitions of life (see, for instance, the definitions from Pályi et al., 2002) cannot satisfy thoroughly the demands of items c) and e), because they deal with only a few attributes of life. As an expressive example, the extremely short definition "Living thing makes models, and nonliving do not" given by Patten et al. (1997) can be cited. It certainly looks rather elegant and ingenious. However, the term "model" can be freely interpreted, so one could say: "My son never made models, is he alive? What nonsense!".

Table 2. Comparison of life's attributes postulated by different authors and the relationship of these attributes to primordial life.

	Ecosystem principle by Patten et al. (1997)	"Pillars of life" by Koshland (2002)	Their possible counter partner in primordial life
1	Conservation	Energy	Retention time for energy of excitement
2	Dissipation	?	Dissipation
3	Openness	?	Openness
4	Growth	?	Irreversible growth
5	Constraint		Initial conditions
6	Differentiation	?	Start of irreversible differentiation
7	Adaptation	Adaptability	Adaptation
8	Coherence		?
9	?	Regeneration	Self-recruitment
10		Program, Seclusion, Improvisation, Compartmentalization	

Widely accepted definitions are also vulnerable. The so-called Darwinian definition goes as follows: "Life is a system capable of evolution by natural selection" (Sagan, 1970); this restricted by only two attributes of life and, consequently, it is unable to differentiate between a population and the life as a global phenomenon. More detailed specification of the Darwinian definition by a supplement "self-sustained chemical system" (Joyce, 1994) does not radically affect the discriminating capacity of the definition, because it is hardly possible to specify life, including early life, before the emergence of the replication machinery (for a review, see Santos et al., 2003).

One could say that the Darwinian definition produces a criterion for the emergence of life because it directly connects the origin of life with the beginning of evolution. That criterion also is not universal since the manifestation of evolution may be "fuzzy". A practical example demonstrating the restricted validity of the definition is the hypothesized extraterrestrial life: How long should we wait for the evidence of its evolution (Fleischaker, 1990)?

Any phenomenon can be defined only within the framework of our knowledge of its nature. Cleland and Chyba (2002) have illustrated this thesis with an example of the definition of water before and after knowledge on the chemical structure of molecule of water. "Water is H_2O" – was their final definition separating water from other substances. In as much as water is only a substance, this definition revealing its structure can be accepted as sufficient. However, life is not a simple substance nor a matter body only. It manifests itself as a process or as a system of very complex hierarchical structures with specific dynamics, diversity of functions, and interrelations with their surroundings (Simon, 2004; Prigogine, 1997, 1999; Vernadsky, 2003).

These manifestations of life are elucidated in literature more or less completely. An historical review of the definitions of life shows that there has been a gradual inclusion of new scientific achievements in the understanding of life. At the turn of the 20th century, dialectical materialism professed that life is a form of the existence of protein bodies supporting themselves by an exchange of matter with their surroundings (our translation from German: "Leben ist die Daseinsweise der Eiweißkörper und diese Daseinsweise besteht wesentlich in der beständigen Erneuerung ihrer chemischen Bestandteile durch Ernährung und Ausscheidung..." Engels, 1948, p. 256). For the sake of the goal of this section, we will concentrate on the term "protein body". Initially, it means "the organism", with the additional characteristic of its "protoplasmic" content. After discovering the leading role of DNA in heredity, there were attempts to improve this definition by substituting the term "protein body" by "DNA body". After the formulation of the RNA-world idea, this definition came into disuse. Nevertheless, it keeps its value for a historical analysis, demonstrating how new knowledge can change the content of the definition of the familiar object.

Life subsystems, such as populations and communities, are included in ecosystems composing the Earth's biosphere (Vernadsky, 2003). Ecosystems, however, include life in its interconnections with its surroundings; and for some ecosystems, their surroundings may be in part alive, namely these interconnections play an important role in our notion that life is inseparable from its environment. This consideration is important both for the definition of life and for the early reconstruction of life. One can state that the life has emerged and continues to exist as an ecosystem.

The ecosystem-defined aspect of current life was actively studied over the last decades in many laboratories. Among others, the works by Jørgensen (1999) and Ulanowicz (2003) with their colleges (Jørgensen et al., 2000; Nielsen and Ulanowicz, 2000) have to be mentioned first of all. Based on these works, the authors identify the main ecosystem attributes. They are: conservation (of energy, matter and information), dissipation, openness, growth, constraint,

differentiation, adaptation, coherence (Patten et al., 1997). The authors believe that all eight characteristics are inseparable for ecosystem operation, however, if some properties are lacking, they can be modified or added later on.

It is of interest to compare these eight basic characteristics with the seven "pillars of life" by Koshland (2002), who listed the main principles for current life (program, improvisation, compartmentalization, energy, regeneration, adaptability, and seclusion). We arranged the comparison in Table 2.

One can see that there is only a limited coincidence between these two lists of life's attributes, and coincidences relate mainly to conservation attributes (ecosystem traits) and adaptability (life and life subsystems [organism, population] traits). However, such attributes as growth, openness and dissipation are also inseparable characteristics of life, including the primordial emerging life, yet they are absent from Koshland's list. The discrepancies in the data in Table 2 seem to be a consequence of the different approaches used: the system approach by Patten et al. (1997) and the structural one by Koshland (2002). At the same time, the choice of life attributes was arbitrary and incomplete in both cases. That reflects the absence of a general consensus as to what attributes of life are really unique and inseparable from life itself. Kompanichenko (2004) had recently calculated that from 230 characteristics of life which were cited in 78 definitions of life in the "Fundamentals of Life" volume (Pályi, 2002), only 19 were in common use. Even in such a case, it seems impossible to include all these characteristics into a short definition of life, but the systemic approach can simplify this task.

3 Basic terms for life definition and the hierarchy of manifestations of life

In publications devoted to the emergence of life, we find many examples of arbitrary manipulations with the terms denoting different levels of the life organization. "Living matter", "organism", "the living" and "life" are often used as equivalents. For example, "life is an emergent attribute of a system whose components assemble an organism" (Zavarsin, 2004, p. 814). From the holistic viewpoint, such a permutation is vulnerable, since it equivalates the emergence of life with the organism level of complexity, whereas life is often defined as the collective feature of organisms (Kauffman, 1993, 2000).

"Living matter" is the only equivalent of "life". It seems reasonable to use this term in a context where life interconnects with surrounding "inert matter". The usage of "living matter" in the sense of a part of an organism is more disputable, since it is very difficult to recognize whether this part is really alive or not. "Organism" and "life" are not equivalents, since any individual organism is only a part of life as a system. An "organism" *per se* is unable to exist for a long time, but it forms part of life as a whole.

Any single organism is only a fragment of the life subsystem, namely, a population which presents a set of related organisms and determines the response of this set to the external effects. "A population", in its turn, is only a part of the life system (or a part of the living moiety of an ecosystem, or a part of a community) since no population alone, isolated from others, can be considered as a natural ecosystem unit. Regarding the term *biota* (combined flora and fauna of a region), Kirby (2002) indicated that "biota" has more a general sense than "organism", but yet not so all-embracing as "biological system", which may include an environment as well. The term biological system, considered in relationship with its environment, can be legitimately considered as a part of an ecosystem.

The term "system" was introduced very early into the definitions of life (see review in Smith, 1986), yet the conceptualization of life as a system is, to date, rather far from completeness. This process was initiated in the middle of the 20th century with the well-known publications of Schrödinger (1945), Bertalanffy (1952), Wiener (1948) and others (for a detailed review see Prigogine, 1997, 1999) to cover an obvious gap between systems obeying the thermodynamic laws and living systems.

Many "thermodynamic" definitions of life have appeared since then. Some of them relate to the origin of life as well. Thus, Schneider and Kay (1994) stated that "life emerges because thermodynamics mandates order from disorder whenever thermodynamic gradients and environmental conditions exist" (p. 171).

Another attempt was reviewed by Luisi (1998), who gave a definition where the term "evolution", as a basic of Darwinism, was not used:

"Life is: *A system which is self-sustaining by utilizing external energy/nutrients owing to its internal process of component production and coupled to the medium via adaptive changes which persist during the time history of the system"*.

It is of great importance that life is defined here as a *self-sustaining* system; it is more correct than *self-reproducible* systems, since *self-reproducible* systems are only the subsystems of life, its agents. Nevertheless, one can notice that, although this and other important attributes of life such as adaptability and utilization of external energy and matter (characteristics of open system) are included, the structure characteristics of life are almost completely lacking in this definition.

The idea of network structure of life's system was developed hereafter in the paper of Ruiz-Mirazo et al. (2004), who stated that:

"'*life'* – in the broad sense of the term – *is a complex collective network made out of self-reproducing autonomous agents whose basic organization is instructed by material records generated through the evolutionary-historical process of that collective network"*.

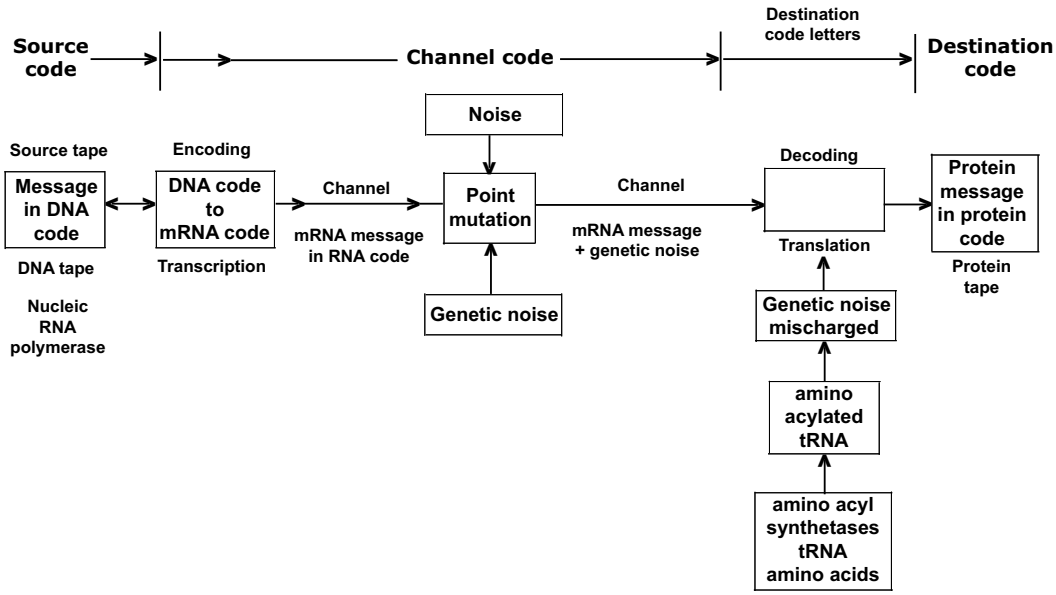

Fig. 1. The transmission of genetic message from the DNA to the protein tape using a transmission channel that is embodied in the gene expression machinery (adopted from Yockey, 2000).

There are two important terms in this definition. The first one, namely *instruction for basic organization*, introduces the informational context into the definition. The second term, e.g. *evolutionary-historical process*, expands our consideration beyond the Darwinian evolution which is usually mentioned as the species oriented evolution (Darwin, 1872; for comments see Erwin, 2000). The development of these important aspects are eagerly anticipated for their contributions in both the recognition of life, in its essence, and the definition of life.

A new trend called the "minimal cell" was initiated in the last decades starting from the works of Varela et al. (1974), Woese (1983), Morowitz (1992) and others (for a review, see Islas, 2004). They try to define a minimal set of genes that can represent the key attributes of life. A minimal set of the protein-encoding genes was shown to be extremely small: a mutant strain of *Mycoplasma genitalium* with 265 to 350 genes can grow and divide under laboratory conditions (Hutchinson et al., 1999). This value is about an order of magnitude lower than the number of enzymes (2000) proposed for the simplest organism about two decades ago (Hoyle and Wickramasinghe, 1978). However, extrapolation of these results to the early evolution of life is rather difficult, since such wild species and their mutants are parasitic species and can exist only under highly favourable conditions. Moreover, those examples represent the reduced variants of organisms living for a long time, so this reduction needs the conditions of operation of the genetic apparatus based on the DNA/RNA/Protein chain of information transfer. Emergence of this apparatus *per se* is one of the core problems of early life.

4 "Biological information", "goal-seeking behavior" . . . , what else?

"...it is absolutely inconceivable to think of the history and major achievements in 20th century biology without the concept of information. In molecular biology one has replication, proofreading, messengers, editing, etc." (Szathmáry, 2001). Nevertheless, most of a recent models dealing with the informational aspects of life were based on Shannon's theory of communication (Shannon, 1948), as well as on information transfer restricted by gene expression. Perhaps the linear character of the gene expression process was a reason why many authors restrict their information quest to canonic informational pathways (see as examples and for references Yockey, 2000; Szathmáry, 2001).

The expression of genetic information includes a transcription of mRNA on a DNA template and the subsequent use of this RNA to govern the synthesis of protein molecule. During transcription/translation events, the linear sequence of DNA nucleotides converts into the amino acid sequence of protein molecule. Thus, transmission of the genetic message is restricted to rewriting the DNA tape information into protein tape information (Yockey, 2000, Fig. 1).

Such a scenario, or similar scenarios can be found in many papers illustrating the applications of information theory in biology. The scheme seems to be elegant (see demand c) of Ruiz-Mirazo et al. (2004), see Sect. 3 of this paper). However, what is noticeable is the surprisingly small value of information of living organisms. Yockey (2000) calculated that an information content of 2000 enzymes carried by one protobiont (the simplest organism of Hoyle and

Wickramasinghe, 1978) amounted to 935 000 bytes. Olson (1995) calculated that the development of a human being is governed by 750 MB of information. Both numbers seem to be incredibly low in comparison with the elementary computer programs we use daily. The cause is to be sought in the nature of the very informational channel and communication itself. In Fig. 1, communication can be compared with a list of samples from some collections proposed, say, for exhibition. Thus, four pictures of a great painter and four pictures in their frames will be coded by an equal numbers of bits. Moreover, for a non-pre-instructed receiver, both communications contain very limited information about the subjects coded. If the receiver is pre-instructed, an additional and supposedly rather significant information value has to be considered. This information partly pre-exists in the system and is partly generated during the decoding process. Decoding starts from the point of identification of an amino acid residue (AA) itself and its position in the protein tape and in relation to its close (molecular, atomic) surroundings. The information content of the latter kind was investigated in the example of protein structure (see, for instance, Nekrasov, 2002, and references therein). There is a great difference between communicatory and decoding mechanisms. The former has nothing to do with the information indicating that this particular AA, for example, arginine, is rather infrequent in natural proteins, and that it possesses a guanidine group and hereafter can participate in different molecular connections, especially with nucleic acids, etc.

The measure of information, that could be extracted from a given symbolic communication, depends on the decoder capability. Although the structures of sender and receiver are absent from Fig. 1, we must emphasize that the whole code supporting machinery is very complex, even much more complex than a structure of channels for the genetic message in this example. This machinery could not be reduced to genetic noise because, in many cases, the organism as a whole can be regarded as a decoder. Emergence of that machinery is a prerequisite of the expression of genetic information, as well as a linear genetic informational carrier. One can conclude that the linear information carriers have had little to do with the evolutionary transition state, since they could appear only during the later stages in the evolution of life.

It is an opportune moment to quantify the term "message". The equation for the measure of information that appears in Shannon (1948) is:

$$H = -K \sum_i p_i \cdot \log p_i \,, \tag{1}$$

where p_i is the probability of the $i-$th member of the alphabet. According to Shannon's explanation, K is a positive constant that "merely amounts to a choice of a unit of measure". K has no physical dimensions and when the logarithm is taken to the base 2 and K=1, H is measured in bits. The informational content of Shannon's communication is small because it is the information of coding symbols and only the symbols *per se*. This is a central affirmation of Shannon's information theory.

Information measured by this Eq. (1) does not mean *knowledge*, it is a mere message composed of a sequence of symbols. In the case of genetic information these symbols are nucleotides and their combinations. However, the Shannon measure of information is still a unique basis of a broad variety of informational aspects related to biological systems, in particular, "knowledge" aspects, "symbolic" aspects, "goal-seeking behavior" (Chernavskii, 2000; Lombardi, 2004).

The expression given in Eq. (1) is very similar to that of entropy in statistical mechanics; for this reason the information was often considered as negative entropy. In the words of Corning (2002): "Negative entropy means, literally, an absence of an absence of thermodynamic order...". Information as an elusive quantity resists definition, although we can determine this quantity and many researchers readily restrict themselves in their search of information to the example of the genetic code. "We are not equipped to attempt any serious navigation of these deep waters", said Patten et al. (1997, p. 234) in a similar situation related to the information content of ecosystem.

We argue also that triplet-amino acid conversion is not the only informational structure with which genetic information operates. For example, the differential activity of genes which underplayed all development programs is determined by the synthesis of some small proteins, known as transcription factors (Fig. 2; for new data on these poly-functional proteins see Rebay et al., 2005). These molecules interact with specific sites on DNA to open a number of informational channels similar to those schematized in Fig. 1. The number of channels can be very large, supposedly, equal to the number of structural genes in DNA; for example, about of 3×10^4 genes in man (Claverie, 2001). Many other ways of controlling information (splicing and alternative splicing, RNA editing, post-transcriptional and post-translation gene silencing and so forth) operate in living organisms, also.

In plants, for example, some significant processes of differentiation started from the synthesis of some molecules of endogenous growth regulators (phytohormones), e.g. auxin (Woodward and Bartel, 2005). The interaction of an auxin molecule with a receptor in a plasma membrane resulted in the "acidic growth" of the cell wall and its surface expansion, which, in turn, induced additional RNA and protein synthesis, DNA duplication and so forth, to finish with the division of a cell. Many small channels may be identified in this way and many other information flows crossing and affecting this one can be listed. The best illustration of the situation can be found in the publications on gene nets (see Kolchanov and Hofestaedt, 2004).

After considering the informational function of life as a net of information flows in an organism, let us look at the information exchange between organisms. It seems to be obvious that the organisms exchange genetic information by the

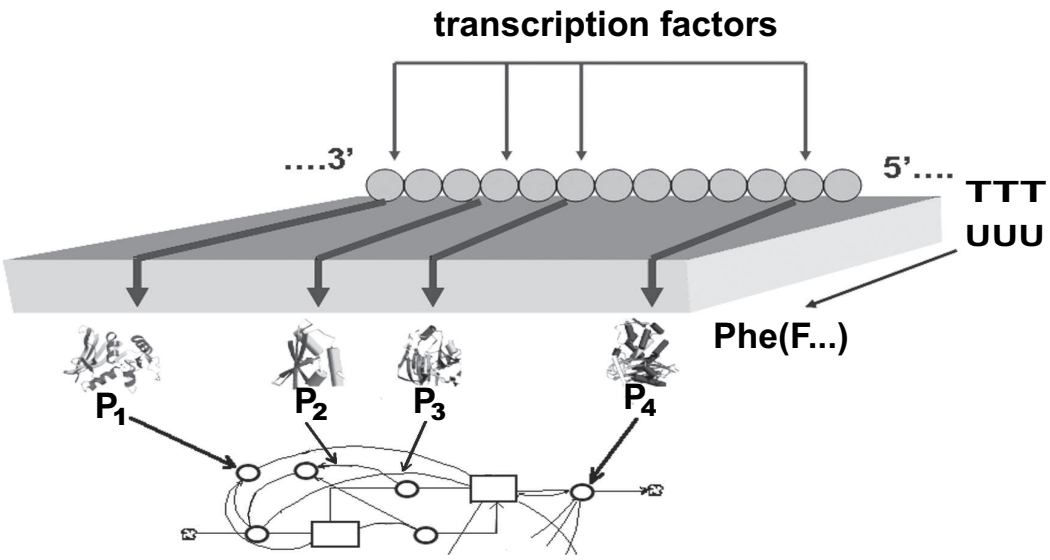

Fig. 2. Scheme of a regulation of gene expression by transcription factors and participation of gene expression products (P1, P2...Pn) in the regulation of metabolic pathways. The scheme illustrates that many information events take place before and after message transmission from DNA tape (TTT...) to protein tape (F...).

sexual process, horizontal gene transfer and other known processes (Lynch, 2002; Arber, 2003; Shapiro, 2002), thus forming the gene pool of populations. The evolutional trajectory of certain genes is defined by natural selection, gene flow or by more neutral processes, such as lineage sorting and others (Avise, 2000; Sites and Marshall, 2003). How much information remains unaccounted for in all of these cases?

There is one more important item we have to emphasize in connection with "*evolutionary-historical process that generates material records*", according to Ruiz-Mirazo et al. (2004). If this process relates to the evolution of life in general, whereas "*material records*" are executed in agents (as only the agents are material and then recordable), we obtain one proclaimed example of an informational exchange between life (sub)systems of different hierarchical levels. Indeed, there is a dramatic gap in our knowledge of information flows between ontogenetic and phylogenetic levels of life. Such a connection, if it exists, might support the evolution of higher taxa, as Gould (2002) hypothesized.

Thus, it is clear that the information flows in living systems are not exhausted with gene expression information. Moreover, the information function of life is not exhausted by the interconnection between subsystems and between different levels of system hierarchy; however, to date, we have no adequate measure of all these informational manifestations of life. In this connection, the comment of one of our anonymous referee sounds very actually: "... the emergence of life could be described as the need of material records (genes) as a way to perpetuate self-maintaining dissipative structures. However, great care has to be used in handling these ideas because genetic information and its biological meanings (ultimately derived from translation) are intimately linked. Moreover, the inheritability of the information coded by genes (responsible for Darwinian evolution) should not be extended to this additional content which is based on the existence of the exchange of information between individual agents and may be attributed to a property of the network rather than that of the agents" (see BG Discussion, 2006, 3, 1–27, referee#1). We quite agree that the informational aspect is one of the most complex and ambiguous ones of all types of life. On the one hand, it seems natural to describe the genetic order in terms of "textual information", and Shannon's measure allows one to evaluate this order. On the other hand, we do clearly realize that living systems are ordered in many different manners, and one often has to create a measure of an order drastically different from the genetic one. Of course, if all types of orders could be reasonably represented as sequences of symbols, say, ones and zeros, Shannon's measure could probably become universally applicable for living systems, too. However, we often find it difficult to do so: it is, for example, the case when we deal with the structural complexity of living systems. We meet a yet more complex case when some sort of a measure is required to evaluate the dynamic complexity of a system, e.g. that of metabolic networks. Use of Shannon's measure is not always well-grounded, although in some cases relevant measures can be constructed using Shannon's formula. On the whole, we do not find any well-defined informational principle (*Urstoff* in German) to be universally applicable for multifaceted contexts of biological order. Therefore, the informational content of life remains so far undefined, and it is a fact we have to face quite independently of the threat it may or may not

represent to the Darwinian paradigm. However, we certainly agree that this subject is yet to be discussed, and we will be happy if our paper could contribute to such a discussion.

We have to supplement also, that in living systems or in a system imitating life, the information function results in the creation of goal-seeking behavior. The goal (in this context) has no measure, nevertheless, it was shown to be very important on all levels of life hierarchy (Bendoricchio and Jørgensen, 1997; Wilhelm and Bruggemann, 2000; Chernavskii, 2000).

Research on the goal function of life, on quantum information theory (see Roy et al., 2002), is now in its early stages and nobody knows how many new attributes of life will be discovered in the future. In our attempts to define life, we are always doomed to balance between incomplete knowledge on the nature of life and our fervent desire to move forward in this field.

5 One more definition of life and conclusion

The most important life manifestations can be combined into three main groups that represent life as: i) a state, ii) a structure, and iii) a process. Consequently, a definition that reflects simultaneously the three-sided view of the life, seems to be comprehensive. However, our knowledge of each of these directions is accomplished to differing extents, and so, such a definition will be neither elegant (as demanded by Emmeche, 1998; Ruiz-Mirazo et al., 2004) nor complete.

Moreover, it is rather difficult to separate one kind of life manifestation from others. Life itself is indecomposable to its constitutes: first of all, the structural and process characteristics are so tightly associated that it seems impossible to describe them in isolation. Nevertheless, we try to give an explicit definition of life, consisting of three parts, to finish with the relationship of this definition to the reconstruction of events leading to the emergence of life:

i) Life, as we see it now, is a specific state of matter (the living state) resulting from the interaction between matter and energy carriers. This interaction starts from the utilization of solar radiation by autotrophic organisms, and spreads over a diversity of organisms via numerous (bio)chemical cycles. A significant part of the utilized energy is retained in organisms by molecular carriers and "network channels" of high energy content; lessening of the utilized energy pool up to some critical level entails in death.

ii) Life on Earth is represented by a specific hierarchical system (the living system) consisting of self-reproducing agents. These agents are the only reference matter of life and are often represented by organisms. They can sometimes be represented as more complex units: bisexual pair, beehive, etc. The agents being individuals can interact with each other and therefore the whole system can be considered as a fragmented and integral entity simultaneously. Different levels of the organization of agents correspond to different levels of life hierarchy. Life as a system shows its worth in the diversity of constraints, feedbacks and interconnections with surroundings.

iii) Life on Earth proceeds as the specific process (the living process). It is expressed in transformations of surroundings (by agents) and in transmutations of the self-reproducing agents themselves. From the physico-chemical point of view, the living process has both dynamic and informational contents. It allows the agents to properly meet the changes in environment and to expand (spread) over a space, thus increasing the level of system complexity and differentiation.

In our reconstruction of the emergence of life we have to keep the three-sided view of life, as specified above. This reconstruction should start from finding such (molecular?) representatives of the self-reproducing agents, which, being the simplest ones, nevertheless, reserve the competence to create the living state, the living system, and the living process.

The first part of this definition is devoted to the characterization of life as a state. We are afraid it is too far from being as complete as the statement "Water is H_2O". The essence of a "specific state" remains obscure, although we suspect that this state is in some way connected with (and may be originated from) an excited state of organic molecules and their ensembles. More detailed knowledge will be obtained from investigations on macromolecular and supramolecular systems and related fields. Thus, we consider the physics and chemistry of complex molecular systems as a more apt tool to clarify the essence of a specific life state and the events leading to its emergence in the transition period. However, in the absence of experimental data we are obliged to restrict this part of the definition by a description of some characteristics we consider as important for understanding of this state.

The second part of the definition characterizes life as a hierarchical system while avoiding the detailed characteristics of the sub-systems and their interrelationships. We should clarify two important aspects of these interrelationships. Firstly, most of known examples of hierarchy embrace the rather long chains of subsystems, starting from genes and the biochemical cycles inside the cell, and continuing with the inner structures of the organism, as well as with communities of organisms to compose the biosphere (McShea, 2001; Oltvai and Barabási (2002). The sub-systems along this line are enclosed in each other, sometimes in a fractal way. Secondly, the hierarchy is represented by a certain (pyramidal or matrix) structure, where its elements reveal both horizontal and vertical connections. The nature of these connections on different levels of hierarchy remains obscure.

With respect to the third part we should emphasize that such accepted attributes of life as the capacity to evolve is not named here, however, it is implied in this part of the definition, ensuing from the set of capabilities of the living process and agents of life, which are listed in this paragraph. The third part of the definition also reflects uncertainties inherent in the question about the leading role of

interconnections between "inert" surroundings and "active" living matter. New molecular findings of gene duplication, DNA rearrangement and DNA acquisition have proved the active position of life in the evolution (Lynch, 2002; Arber, 2003; Shapiro, 2002). According to this knowledge, agents of life are transmuted in such a way to produce a diversity of genovariants among which the more fitted will be selected during future changes in the environment. However, nobody knows whether the environmental changes are accidental or whether they follow a certain scenario, predetermining a direction in the development of life.

The basic attributes of open systems – energy and matter flows – are hidden in the expression "*standing in the multifold interconnections with surroundings*". We were forced to limit ourselves in this field due to the existence of many controversial opinions with regard to the first energy source used by primordial life; each energy source presupposes a unique way of transition from chemical to biological evolution. "The flow of energy through a system acts to organize that system", said Harold Morowitz (1968). There is a huge body of literature on this topic which cannot be reviewed here and a such review, if undertaken, will be unable to bring us to broad acceptable conclusions.

Regrettably, the information essence of life is not displayed in our definition of life. This is a consequence of the incompleted knowledge of the nature of information itself. "Different combinations of energy and matter arrayed in space or sequenced in time give rise in the elusive quantity, information" (Patten et al., 1997, p. 234). Moreover, information theory at present reveals its structural complexity; at least three concepts of information have been recently formulated (Lombardi, 2004).

The third part of the definition shows that the splitting (perhaps only in one's mind) of events of the transition into two moiety relating "agents" and "system" respectively, simplifies the reconstruction of the transition state of evolution. This is because certain life attributes, which appear with difficulty in agents, can easily emerge in the system.

Finally, it becomes more and more clear that the transition period has to be short, but more saturated with the events that were fundamental to the conversion of the non-living compositions into living systems.

In our work we do not compared our vision of life and evolution with the other publications in this field (e.g. most of contributors of OLEB proposed their own scenarios). Certainly, the definition proposed in this paper may be considered as a validity test for these scenarios. However, we have wittingly tried to avoid such analysis, first of all because this paper represents only a definitive part of our vision. To treat the problem comprehensively, we are going to discuss this question in our next publications, where we intend to consider life as a structure, a state, and as a complex dynamic process.

Acknowledgements. This work was partly supported by the Program of Presidium of RAS "Origin and evolution of biosphere", project no. 04-1-II25-035 FEB RAS, the Program COST D-27 "Prebiotic Chemistry and Early Evolution" of the European Science Foundation, and RFBR-05-03-32563.

Authors thank two anonymous referees for their valuable comments and recommendations to continue this work; we also are greatly appreciated to the handling editor F. Westall who promoted clarification of some principal thesis of this paper. Y. Zhuravlev is grateful to E. Y. Frisman for his helpful discussion of these issues which partly resulted from the joint work on the dynamics of early non-linear living systems, to L. Hlopkova who greatly improved the English of the paper, and to E. Sundukova for her excellent assistance on the all stages of manuscript preparation.

Edited by: F. Westall

References

Arber, W.: Elements for a theory of molecular evolution, Gene, 317, 3–11, 2003.

Avise, J. C.: Phylogeography. The History and Formation of Species, Harvard University Press, Cambridge/Massachusetts/London, England, 2000.

Bendoricchio, G. and Jørgensen, S. E.: Exergy as goal function of ecosystems dynamic, Ecological Modelling, 102, 5–15, 1997.

Bertalanffy, von, L.: Problems of Life: an Evaluation of Modern Biological Thought, J. Wiley, N.Y., 1952[1949].

Bohr, N.: Light and life, Nature, 308, 456–459, 1933.

Chernavskii, D. S.: The origin of life and thinking from the modern physics point of view (in Russian), Uspechi Physich, Nauk, 170, 157–183, 2000.

Claverie, J.-M.: What if there are only 30 000 human genes?, Science, 291, 1255–1257, 2001.

Cleland, C. E. and Chiba C.: Defining "life", Origins Life Evol. Biosph., 32, 387–393, 2002.

Corning, P. A.: Thermoeconomics: Beyond the second law, J. Bioeconomics, 4, 57–88, 2002.

Darwin, C.: The Origin of Species, VI Ed., Studio Editions Ltd., Princess House, London, 1872.

Eigen, M. and Schuster, P.: The Hypercycle: a Principle of Natural Self-organization. Berlin: Springer-Verlag, 1979.

Eigen, M., Schuster, P., Gardiner, W., and Winkler-Oswatitsch, R.: The origin of genetic information, Sci. Am., 244, 78–94, 1981.

Emmeche, C.: Defining life as a semiotic phenomenon, Cybernet. Human Knowing, 5, 3–17, 1998.

Engels, F.: Dialectics of Nature, OGIS, (Russian translation), Moscow, 1948.

Fleischaker, G. R.: Origins of life: an operational definition, Origins Life Evol. Biosph. 20, 127–137, 1990.

Erwin, D. H.: Macroevolution is more than repeated rounds of microevolution. Evolution and Development, 2, 78–84, 2000.

Gesteland, R. F., Cech, T. R., and Atkins, J. F. (Eds.): The RNA World, 2nd ed., Cold Spring Harbor Press, 1999.

Gilbert, W.: The RNA world, Nature, 319, 618, 1986.

Gould, S. G.: The Structure of Evolutionary Theory, Belknap, Harvard Univ. Press, Cambridge, MA, 2002.

Haldane, J. B. S.: The origin of life, Rationalist Annu., 148–169, 1929.

Hoyle, F. and Wickramasinghe, N. C.: Life Cloud: The Origin of Life in the Universe. Harper and Row, NewYork, 1978.

Hutchinson, C. A., Peterson, S. N., Gill, S. R., Cline, R. T., White, O., Fraser, C. M., Smith, H. O., and Venter, J. C.: Global transposon mutagenesis and a minimal mycoplasma genome, Science, 286, 2165–2169, 1999.

Islas, S., Becerra, A., Luisi, P. L., and Lazcano, A.: Comparative genomics and the gene complement of a minimal cell, Origins Life Evol. Biosph., 34, 1–2, 243–256, 2004.

Jørgensen, S. E., Patten, B. C., and Straskraba, M.: Ecosystems emerging: 4. growth, Ecological Modelling, 126, 249–284, 2000.

Jørgensen, S. E.: State-of-the-art of ecological modelling with emphasis on development of structural dynamic models, Ecological Modelling, 120, 2–3, 75–96, 1999.

Joyce, G. F.: in: Origins of Life: the Central Concepts, edited by: Deamer, D. W., Fleischaker, G. R., Jones and Bartlett, Boston, p. xi–xii, 1994.

Kauffman, S.: Investigations, Oxford University Press, Oxford, Ch. 3, 2000.

Kauffman, S.: The Origins of Order: Self-organization and Selection in Evolution, Oxford, University Press, Oxford, 1993.

Kirby, K. G.: Biological adaptabilities and quantum entropies, BioSystems, 64, 33–41, 2002.

Kolchanov, N. and Hofestaedt, R. (Eds.): Bioinformatics of Genome Regulation and Structure, Kluwer Academic Publishers, Boston/Dordrecht/London, 2004.

Kompanichenko, V.: Systemic approach to the origin of life, Frontier Perspectieves, 13, 22–40, 2004.

Koshland Jr., D. E.: The seven pillars of life, Science, 295, 2215–2216, 2002.

Lombardi, O.: What is information?, Foundation of Science, 9, 105–134, 2004.

Luisi, P. L.: About various definitions of life, Origins Life Evol. Biosph., 28, 4–6, 613–622, 1998.

Lynch, M.: Gene duplication and evolution, Science, 297, 945–947, 2002.

McShea D. W.: The minor transitions in hierarchical evolution and the question of a directional bias, J. Evol. Biol., 14, 502–518, 2001.

Morowitz, H. J.: The Beginnings of Cellular Life, Yale University Press, New Haven, 1992.

Morowitz, H. J.: Energy flow in biology. Academic Press, New York, 1968.

Nekrasov, A. N.: Entropy of protein sequences: an integral approach, J. Biomol. Struct. Dyn., 20, 87–92, 2002.

Nielsen, S. N. and Ulanowicz, R. E.: On the consistency between thermodynamical and network approaches to ecosystems, Ecological Modelling, 132, 23–31, 2000.

Olson, M. V.: A time to sequence, Science, 270, 394–396, 1995.

Oltavai, Z. N. and Barabási, A.-L.: Life's complexity pyramid, Science, 298, 763–764, 2002.

Oparin, A. I.: Proiskhozhdenie zhizny (in Russian), Izd. Mosk. Rabochii, Moscow, 1924.

Orgel, L. E.: Some consequences of the RNA world hypothesis, Origins Life Evol. Biosph., 33, 211–218, 2003.

Pályi, G., Zucci, C., and Caglioti, L. (Eds.): Fundamentals of Life. Elsevier, Paris, 2002.

Patten, B. C., Straškraba, M., and Jørgensen, S. E.: Ecosystems emerging: 1. conservation, Ecological Modelling, 96, 221–284, 1997.

Prigogine, I.: Nonlinear science and the laws of nature, Int. J. Bifurcat. Chaos, 7, 1917–1926, 1997.

Prigogine, I.: Laws of nature, probability and time symmetry breaking, Physica A, 263, 528–539, 1999.

Rebay, I., Silver, S. J., and Tootle, T. L.: New vision from eyes absent: transcription factors as enzymes, Trends in Genetics, 21, 163–171, 2005.

Roy, P. K., Miller, J. P., and Majumder, D. D.: A control analysis of neuronal information processing: a study of electrophysiological experimentation and non-equilibrium information theory, Lecture Notes in Computer Science, 2275, 191–203, 2002.

Ruiz-Mirazo, K., Peretó, J., and Moreno, A.: A universal definition of life: autonomy and open-ended evolution, Origins Life Evol. Biosph., 34, 323–346, 2004.

Sagan, C.: Life, in: The Encyclopaedia Britannica, William Benton, London, 1970.

Santos, M., Zintzaras, E., and Szathmáry, E.: Origin of sex revisited, Origins Life Evol. Biosph., 33, 405–432, 2003.

Schneider, E. D. and Kay, J. J.: Life as a manifestation of the second law of thermodynamics, Mathematical Computer Modelling, 19, 25–48, 1994.

Schrödinger, E.: What Is Life?, Cambridge Univ. Press, 1945.

Segre, D., Ben-Eli, D., Deamer, D. W., and Lancet, D.: The lipid world, Origins Life Evol. Biosph., 31, 119–145, 2001.

Shannon, C. E.: A mathematical theory of communication, Bell Syst. Tech. J., 27, 379–424, 623–656, 1948.

Shapiro, J. A.: A 21st century view of evolution, J. Biol. Phys., 28, 745–764, 2002.

Simon, H. The Sciences of the Artificial (Russian translation), Editorial URSS, Moscow, 2004.

Sites, J. W. and Marshall, J. C.: Delimiting species: a Reneissance issue in systematic biology, Trends in Ecology and Evolution, 18, 462–470, 2003.

Smith, M. J.: The Problems of Biology, Oxford University Press, Oxford, 1986.

Szathmáry, E.: Biological information, kin selection, and evolutionary transitions, Theor. Popul. Biol. 59, 11–14, 2001.

Szathmáry, E. and Demeter, L.: Group selection of early replicators and the origin of life, J. Theor. Biol., 128, 463–486, 1987.

Ulanowicz, R. E.: Some steps toward a central theory of ecosystem dynamics, Comput.Biol Chem., 27, 523–530, 2003.

Varela, F. J., Maturana, H. R., and Uribe, R.: Autopoiesis: the organization of living systems, its characterization, and a model, Curr. Mod. Biol., 5, 187–196, 1974.

Vernadsky, V. I.: Biosphere and Noosphere (Russian edition), Airis-Press, Moscow, 2003.

von Bertalanffy, L.: Problems of Life: an Evaluation of Modern Biological Thought, J. Wiley, N. Y., 1952 (1949).

Wächtershauser, G.: Before enzymes and templates: theory of surface metabolism, Microbiol. Rev., 52, 452–484, 1988.

Wiener, N.: Cybernetics: or Control and Communications in the Animal and the Machine. MIT Press, Cambridge, MA, 1948.

Wilhelm, T. and Bruggemann, R.: Goal functions for the development of natural systems, Ecological Modelling, 132, 231–246, 2000.

Woese, C. R.: The primary lines of descent and the universal ancestor, in: Evolution from Molecules to Man, edited by: Bendall, D. S., Cambridge University Press, Cambridge, 209–233, 1983.

Woodward, A. W. and Bartel, B.: Auxin: regulation, action, and interaction, Annals of Botany (London), 95, 707–735, 2005.

Yockey, H. P.: Origin of life on earth and Shannon's theory of communication, Computers and Chemistry, 24, 105–123, 2000.

Zavarzin, G. A.: The future is selected by the past (in Russian), Vestnik RAN, 74, 813–822, 2004.