Origins of Life and Evolution of Biospheres

Volume 40 • Numbers 4–5 • October 2010

Special Issue: Workshop OQOL'09: OPEN QUESTIONS ON THE ORIGINS OF LIFE 2009
Guest Editors: Kepa Ruiz-Mirazo • Pier Luigi Luisi

Further articles can be found at www.springerlink.com

Abstracted/Indexed in: *Academic OneFile, AGRICOLA, Astrophysics Data System (ADS), Biological Abstracts, BIOSIS Previews, Chemical Abstracts Service (CAS), CSA/Proquest, Current Awareness in Biological Sciences (CABS), Current Contents/ Life Sciences, Elsevier Biobase, EMBASE, EMBiology, Environment Index, Gale, GeoArchive, Geobase, GeoRef, Geotitles, Google Scholar, IBIDS, Index Copernicus, Journal Citation Reports/Science Edition, OCLC, PubMed/Medline, Science Citation Index, Science Citation Index Expanded (SciSearch), SCOPUS, Summon by Serial Solutions*.

Instructions for Authors for *Orig Life Evol Biosph* are available at http://www.springer.com/11084.

The subject of the origin and early evolution of life is an inseparable part of the general discipline of Astrobiology. The journal *Origins of Life and Evolution of Biospheres* places special importance on the interconnection as well as the interdisciplinary nature of these fields, as is reflected in its subject coverage. While any scientific study which contributes to our understanding of the origins, evolution and distribution of life in the Universe is suitable for inclusion in the journal, some examples of important areas of interest are: prebiotic chemistry and the nature of Earth's early environment, self-replicating and self-organizing systems, the theory of the RNA world and of other possible precursor systems, and the problem of the origin of the genetic code. Early evolution of life–as revealed by such techniques as the elucidation of biochemical pathways, molecular phylogeny, the study of Precambrian sediments and fossils and of major innovations in microbial evolution–forms a second focus. As a larger and more general context for these areas, Astrobiology refers to the origin and evolution of life in a cosmic setting, and includes interstellar chemistry, planetary atmospheres and habitable zones, the organic chemistry of comets, meteorites, asteroids and other small bodies, biological adaptation to extreme environments, life detection and related areas. Experimental papers, theoretical articles and authoritative literature reviews are all appropriate forms of submission to the journal. In the coming years, Astrobiology will play an even greater role in defining the journal's coverage and keeping *Origins of Life and Evolution of Biospheres* well-placed in this growing interdisciplinary field.

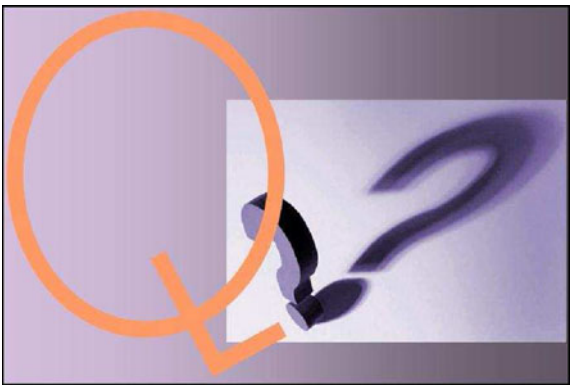
Origins of Life and Evolution of Biospheres is published bimonthly.
Periodicals postage paid at Rahway, N.J. USPS No. 491–630.
U.S. Mailing Agent: Mercury Airfreight International Ltd., 365 Blair Road, Avenel, NJ 07001 Published by Springer, P.O. Box 17, 3300 AA Dordrecht, The Netherlands, and 101 Philip Drive, Norwell, MA 02061, U.S.A.
Postmaster: Please send all address corrections to: *Origins of Life and Evolution of Biospheres*, c/o Mercury Airfreight International Ltd., 365 Blair Road, Avenel, NJ 07001, U.S.A.

printed on acid-free paper

Workshop OQOL'09: OPEN QUESTIONS ON THE ORIGINS OF LIFE 2009

Palacio Miramar, SAN SEBASTIAN - DONOSTIA, SPAIN, MAY 20-23, 2009

Extended Abstracts



Springer

Open Questions on the Origins of Life: Introduction to the Special Issue

The origin of life on Earth is still a mystery, one of the greatest mysteries in science today. We are surrounded by myriads of life forms—each leaf of a tree in a forest contains billions of living cells, our body contains huge numbers of active microorganisms, we people keep living and growing, incessantly—and we do not yet know how life came about on our planet. Our ignorance about the origin of life is profound—not just some simple missing mechanistic detail. We do not know how the genetic code came about, we do not understand yet how the specific sequences of proteins or nucleic acids came about in multiple identical copies, we do not have a precise idea about the structure and functioning of the first protocells. This ignorance stems not only from our experimental difficulties with prebiotic chemistry, but is also conceptual, as we are not yet able to conceive on paper how all these things came about.

There are two classes of people who look at things with a much more positive perspective. One group includes the staunch adherents of the RNA-world: give them a family of self-reproducing ribozymes and all is solved, at least in theory. Nevertheless, one of the main problems is precisely this, that nobody is able to conceive the formation of sophisticated prebiotically self-constructed RNA replicases. Where did that RNA come from? We all hope, of course, that a satisfactory answer will be given in the very near future, and this will be a great day for the field of the origin of life. For the moment, however, the "prebiotic RNA-world" remains a pious dream.

The other optimist front, although in a different context, consists of those people who believe firmly that the origin of life on Earth was an obligatory pathway—it had to happen, it could not have been otherwise. A key representative of this way of thinking is Christian de Duve—e.g.: see his last (2009) paper on the subject of contingency and determinism. This view is tantamount to saying that the initial conditions on our planet were such, that the occurrence of life was an event of very high probability—"the gospel of inevitability", as Szathmary calls it.

However, precisely the fact that we, chemists and biologists, and scientists at large, after more than 50 years of intelligent effort, do not see any way of making life in the laboratory should be a clear demonstration that life does not form so easily and spontaneously. Otherwise, we would have found it by now. And the reason why this is not so is quite clear, too: life formation is not a process under thermodynamic control, running towards a minimum of energy. It is made possible by catalysts, disseminated in the path just to avoid the downhill thermodynamic flow of reactions. Under this perspective, the idea that the formation of life on Earth is a spontaneous, easy process, which had to occur *sic et simpliciter*, appears rather extravagant.

Actually, there are open questions about the origin of life just because this is, indeed, a difficult problem. That is the basic motivation for the workshop we organized, out of which the following collection of extended abstracts has been derived.

So the overall idea behind the OQOL workshop is to ask which are the main stumbling blocks in this pathway of discovery, why are they so difficult to solve, and also to possibly shed light on what we should do to make some real progress in the near future.

The original structure of the meeting was to formulate and offer a number of these questions in advance to the contributors, asking them to choose the one(s) they would like to tackle, provided they do so as directly as possible, i.e., without digressing too

much in their own standard talk. A similar meeting was already held, in a preliminary form, in Erice, Sicily, in 2006, and raised quite a bit of interest. Many researchers asked to continue the experiment, and thus this second edition in San Sebastian came along, 3 years later.

In order to define the eight questions which finally were debated at the meeting in 2009, about eighty researchers were contacted by e-mail and asked to formulate which were for them the most important open questions in the field. Out of this first sampling, fourteen questions were selected by our committee and, in a second round, we asked the fifty-sixty researchers who responded positively to select eight out of these fourteen. Not a bad example of scientific democracy, so we are grateful to all those who got involved in the process, even if part of them could not attend and contribute to the actual meeting, months later.

The final list of questions was:

1. To what extent the origins of life were deterministic or based on contingency?
2. Is life an emergent property?
3. Was the origin of life heterotrophic or autotrophic?
4. What were the origins of catalytic cycles?
5. How plausible is the "RNA world" hypothesis?
6. How to bridge a gap between the protocellular world and the minimal cell?
7. Is life a unity or confederacy?
8. How to define the very origin of life?

This procedure, interestingly, also gives a rough idea of the climate of the field itself. Note, for example, that the first two questions are philosophical in nature; that some of them have been with us from the very beginning of the field, while some others (in their formulation, at least) are totally new.... In addition, the result of the call, in terms of the number of extended abstracts submitted for each of the questions is a good indicator, as well. This time there were two clear "winners" in that sense: the question on the 'RNA-world' and the one on the 'protocellular world//minimal cell' (questions number 5 and 6 in the previous list).

The meeting consisted of 8 sessions spread over 4 days (one session in the morning and one in the afternoon). Each session focused on a single question, which was addressed by three to four panelists (the speakers) in brief (20 min) presentations. The presentations were then followed by an open general discussion. Thus, each session lasted approximately 3.5 h. This format was chosen because it certainly helps to involve all participants in the debates, including the younger researchers.

The collection of extended abstracts being published in the present special issue of Origins of Life and Evolution of Biospheres gathers not only a summary of the contributions of the panelists to each of the questions they addressed but also, in several cases, their views on some other question that motivated their thinking, plus a few additional selected texts from researchers who attended the meeting but—due to time constraints—did not have the chance to speak as panelists (although they actively took part in the general discussions). To all of them, contributors and participants, our most sincere thanks, because they made it happen, they kept a very high standard (both in terms of the academic level of the exchanges and the attitude during the debates) and they are, therefore, the ones ultimately responsible for the end result, as it appears in this special issue.

Finally, we, as organizers (but surely on behalf of all participants), would like to congratulate Ada Yonath, a very active contributor during our meeting, for winning the Noble Prize in Chemistry, last October, 2009.

For more details about the workshop, please, see: http://oqol2009.wordpress.com/

It is the hope of all of us that the initiative "Open questions on the origin of life" will continue after the first two milestones of Erice and San Sebastian. Each of you is asked to take the lead. And good luck.

Pier Luigi Luisi and Kepa Ruiz-Mirazo
Department of Logic & Philosophy of Science - Biophysics Research Unit (CSIC - UPV/EHU)
UNIVERSITY OF THE BASQUE COUNTRY
E-mail: luisi@mat.ethz.ch; kepa.ruiz-mirazo@ehu.es

February, 2010.

Acknowledgement
The organization of our workshop OQOL'09 in San Sebastian–Donostia, would have been impossible without the financial support from various local institutions, in particular: The Basque Government, The Spanish Ministry for Science (MICINN), The University of the Basque Country (Research, Campus and Language Vice-President Bureaus) and the Spanish Research Council (CSIC). The Department of Logic and Philosophy of Science (UPV/EHU) and the Biophysics Research Unit (CSIC-UPV/EHU) should also be acknowledged for organizational support, together, of course, with the *IAS Research Group*, whose members were the core of the Local Organizing Committee. Finally, we are grateful to Alan Schwartz, again, for his encouragement to publish this collection of extended abstracts in Origins of Life and Evolution of Biospheres and for the editorial support we got from him in order to meet that end.

Workshop OQOL'09

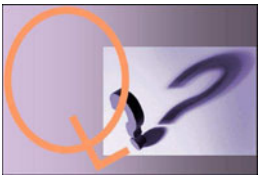
Extended Abstracts for the Following Selected Question

- **Contingency Versus Determinism in the Origin of Life/Origin of Proteins**

Premise. The origin of life is often seen in terms of two basic, opposite schemes, determinism and contingency. Generally, the two principles work hand in the hand, as each "choice" made by contingency must then comply to the natural laws and, in turn, contingency arises from a given thermodynamic asset. However, when we ask the basic question of whether the origin of life follows an obligatory deterministic pathway (absolute determinism), or whether it is due to the vagaries of contingency, the two views become again drastically opposite to each other. More precisely, according to the deterministic view (as represented most notably by Christian de Duve), the origin of life is seen as an event of very high probability: actually, it had to come out inevitably from the starting and boundary conditions (the so called "gospel of inevitability"). The opposite view (advocated, for example, by Jacques Monod), implies that the origin of life was due to the occurrence of several independent factors, each of them perhaps not un-deterministic, whose simultaneous and unpredictable interaction led to successive events, up to the origin of life.

The question. Do you agree that the choice between these two extreme points of view cannot be done on a rational, scientific basis, and is instead for each scientist a matter of philosophical or religious belief? And, if you do not agree, which

scientific arguments would you offer in favour of one or the other lines of thought?



Contingency and Determinism in the Origin of Life-and Elsewhere

Pier Luigi Luisi

Department of Biology, University of Roma Tre, Rome, Italy

Keywords Contingency and Determinism: Contigency • Determinism • Causality

About the origin of life on Earth we generally accept the Oparin-Haldane scenario, according to which life was formed from inanimate matter throughout a long series of spontaneous steps of increasing molecular complexity, up to the formation of the first self-reproducing protocells. From this scenario, any transcendent act or miracle intervention is eliminated by definition.

But how did this series of steps come about? One way to give an answer is in terms of determinism, according to which the laws of physics and chemistry determine sequentially and causally the obliged series of events. Thus, in his book about the origin of life, Christian de Duve (1991) writes:

"…Given the suitable initial conditions, the emergence of life is highly probable and governed by the laws of chemistry and physics…"

This seems to lead one to the idea that life on Earth was inescapable and, in fact, De Duve, in a more recent work (2002), re-states this concept:

"…It is self-evident that the universe was pregnant with life and the biosphere with man"

The idea of the inevitability of life on Earth, although phrased differently and generally with less emphasis, is presented by some other significant authors. For example Harold Morowitz, in his well known book (1992), states:

.."We have no reason to believe that biogenesis was not a series of chemical events subject to all of the laws governing atoms and their interactions" (p.12)

adding also, interestingly:

" Only if we assume that life began by deterministic processes on the planet are we fully able to pursue the understanding of life' origins within the constraints of normative science"(p.3)

plus a clear plead against contingency:

"..We also reject the suggestions of Monod that the origin requires a series of highly improbable events and cannot be recovered from the laws of physics"(p.13)

The interesting conjunction of de Duve's and Morowitz's view is the rejection of the miracle scenario, and the acceptance however of the notion of the inevitability of life *via* the deterministic laws of physics and chemistry.

The "gospel of the inevitability" of the origin of life on Earth, as Eörs Shatmary properly calls it (2002), has its counterpart in the notion that contingency is the basic creative force for shaping the molecular and evolutionary constructs on Earth. This view is not new—actually is an old icon of the history of science. In our contemporary scientific era, one may recall Francois Monod with his "Chance and Necessity" (1971), his colleague Francois Jacob with "The possible and the actual" (1982); and the many books by Stephen J. Gould (see, for example: 1989), who is perhaps the most cited author about contingency in biological evolution.

Contingency may be defined as the outcome of a particular set of simultaneous concomitant effects that apply in a particular point of time/space. In most of the epistemological literature this word has aptly taken the place of the term "chance" or "random event"—and in fact it has a different texture. For example a car accident can be seen as a chance event, but indeed it is due to the concomitance of many independent factors, like the car speed, the road conditions, the state of the tires, the level of alcohol of the driver, etc. These factors all conjure together to let out the final result, seen as a chance event. The same can be said for a stock market crush, or the stormy weather of a particular summer day. Interestingly, each of these independent factors can be seen per se' as a deterministic event, e. g. the bad state of the car tires determines per se' a car sliding off at a curve. The fact however that there are so many of these factors acting simultaneously, and each with an unknown statistical weight, renders the accident unpredictable—a chance event. Change the contingent conditions-perhaps only one of them—and the final result would be quite different-it may happen 1 week later, or with another driver—or never. If it would start all over again in the history of biological evolution—says Stephen J. Gould (1989)

> *"..run the tape again, and the first step from prokaryotic to eukaryotic cell may take 12 billions years instead of 2…"*

implying that the onset of multicellular organisms, including mankind, may have not arisen yet-or may never arise. This is contingency in the clearest form. The most characteristic feature of contingency is seen in the statement "it could have not happened". This is really in contrast to the deterministic view "it must happen".

I would like to examine the old dichotomy between determinism and contingency by using a concrete case in the field of the origin of life, namely the synthesis of functional macromolecules, such as enzymes and RNA. How these specific sequences—at least 100 residues long—were formed in the prebiotic Earth, in many identical copies, we do not know. One thing we know for sure, however: that lysozyme is not with us because it is more stable than its billions of constitutional isomers. And the same can be said for t-RNA^{phe} and for all the other biologically active macromolecular sequences on Earth. In this regard, the numerology in the field of macromolecular genesis is well known, perhaps somewhat abused. Nevertheless, let me repeat one old calculation. Focussing on a co-polypetide chain of length 60, and having at disposal the 20 classic aminoacids, the number of different chains that are, in principle, possible is 20^{60}, or ca. 10^{70}. The ratio between these two numbers (the possible and the actual, borrowing an expression from Jacob) corresponds approximately to the ratio between the space of the entire universe with respect to the space occupied by one atom (actually, a fraction of this).

How and why these few proteins which are with us have been chosen? Determinism, or contingency?

Determinism would imply, in chemical terms, the idea that the extant proteins (or nucleic acids) are with us because they have something special from the thermodynamic point of view, for example thermodynamic stability, or solubility, or folding, or…. Are there unknown selection

rules that have operated during the prebiotic molecular evolution, choosing the sequence of lysozyme over all others by unknown criteria of kinetic control? This cannot be excluded. But at the present stage one should accept the view that these few proteins of life are with us as the products of the bizarre laws of contingency, followed by chemical evolution processes.

Does this mean that the formation of macromolecules on Earth is a process outside the laws of chemistry and physics? Of course no. Statistical processes belong to science, have only the bad taste of giving an outcome unpredictable in the details.

And the chemistry of macromolecules formation is only one of the many chemistry examples of contingency that can be given in the field of the origin of life. The selection of homochirality, the selection of one particular type of mononucleotide containing ribose and phosphate, the process of formation of the genetic code; the setting of a particular metabolism rather than another one; and the assembly of a full fledged genomic cell—appear all processes where contingency, guided and attended by molecular evolution, played most likely the fundamental creative role.

And we have other dramatic examples of the importance of contingency in the history of life on earth. The fall of a meteorite 65 million years ago destroyed all dinosaurs, which had reigned on our planet for over 150 million years. This impact was clearly the effect of chance—and without that, most probably our planet would be inhabited by them, and not by mankind.

Or we could also mention the invention of oxygen 2 billion years after the origin of life—a random mutation, which had the effect of creating the forms of life which now reign on the earth. Without that random mutation, things on our Earth would have been quite different, and again perhaps no mankind.

The view of contingency, in particular that mankind is the product of "chance", is not very appealing from the emotional point of view. This has been seen clearly by Monod (1971), with his famous notion of "being alone in the universe".

The large field of contingency in biological evolution is outside the frame of this short assay. It is important however for me to state that this "being alone in the Universe" should not lead one to deduce that the humanistic and ethical values are deprived of meaning—that the "sacredness" of life, if you want to call it in this way—is impoverished. I believe in the contrary, that namely the values of consciousness and ethics can be arrived at from within the human construction, without the need of being imported from transcendental sources.

Coming back to the more modest framework of macromolecular chemistry, the statement about the importance of contingency may appear to many very trivial. On the other hand, as we have seen, there is a significant part of the scientific community who rather believe in the inevitability of life out of a deterministic pathway. Is the aim of this assay to point out this contradiction in the present generation of life sciences; and actually, in addition to authors within the origin of life field of research, that there are other cultural movements that, *mutatis mutandi*, echo this anti-contingency view.

One is the anthropic principle. This can be expressed in different forms, but the basic idea is that the universal constants, the geometric parameters and all constants of the universe are the way they are in order for life and evolution to develop. For example, one reads in Paul Davies' (1999) book:

> *"If life follows from (primordial) soup with causal dependability, the laws of nature encode a hidden subtext, a cosmic imperative, which tell them: Make life! And, through life, its by-products, mind, knowing, understanding…".*

And even men above suspicion, like Freeman Dyson, have their.. "it is as if"(1979):

> *"As we look out in the universe and identify the many accident of physics and astronomy that have worked together to our benefit, it almost seems as if the Universe must in some sense have known that we were coming".*

Or a citation by Stephen Hawking (Hawking & Penrose 1996):

> *"And why is the universe so close to the dividing line between collapsing again and expanding indefinitely?…If the rate of expansion one second after the BigBang had been less by one part in 10^{10}, the universe would have collapsed after a few million years. If it had been greater by one part in 10^{10}, the universe would have been essentially empty after a few million years. In neither case would it have lasted long enough for life to develop. Thus one either has to appeal to the anthropic principle or find some physical explanation of why the universe is the way it is."*

I ascribe the anthropic principle to the general category of absolute determinism, as it corresponds to the notion of the "inescapability of life"—which for me contains implicitly the belief in the divine creation or at least of an intelligent design. If the things in the universe are the way they are so that mankind could arise and evolve, then somebody must have conceived this intelligent design.

Let us now consider another scientific movement that—so it seems—operates outside the framework of contingency. This is the field of SETI (Search for Extra Terrestrial Intelligence): i.e., scientists who are trying to catch signals from the cosmos, believing that there is a finite probability that alien civilizations exist and willing to communicate with us (Tarter 2001; Carr 2001). The number of people—including scientists—who believe in this is very large. To that, a personal note: I was recently at a meeting in Windsor, UK, with a small group of scientists discussing questions related to origin of life and physical and chemical tuning in evolution. Freeman Dyson gave a lively after-dinner lecture, describing a complex mirror that would help getting intelligent signals from the outer space. To my question at the end, whether he would really believe in extra-terrestrial intelligence, he answered with a decided no, which set a wave of protest in the small audience. Following this, we decided to take a half-serious poll: how many would say yes, that they believe in the possibility of ETI, and how many would say no, namely be highly sceptical about the possibility of existence of alien civilizations. Twenty-five out of the thirty people said yes. Although the audience was pre-selected and biased towards SETI and anthropic principle, the large defeat of the contingency flag was unexpected and somewhat extraordinary.

With SETI, the contrast with those who stand behind contingency could not be greater. The assumption of intelligent life similar to our's on another planet is based on the unproven assumption that the same set of conditions are operative on that other planet. Not only should one then believe in the determinism of life on our planet, but also on a kind of cosmic determinism that leads to the occurrence of life on other planets. Determinism squared.

Again, it is far from the spirit of this article to throw a spear against SETI. Personally, I think that this is a great vision, and that visions in science should be encouraged—particularly in an era in which mostly pragmatic and applied research projects find support. The point in this paper is rather to emphasize that also this movement is based on the same set of beliefs, on the common faith that life is inevitable and, as such, widespread.

Conceptually close to the idea of SETI is the idea of a general panspermia, which assumes that life on Earth has been originated elsewhere in the universe and came to us in form of some non well identified germs of life. In the more general and poetic version, the theory of panspermia foresees life as a general property that permeates the cosmos and that, therefore, does not need to have an origin (Mastrapa et al. 2001; Hoyle and Wickramasinghe 1999).

Generally then, from the one hand there is in our life sciences the tendency to accept the rationality of contingency-on the other one there is the tendency to reject it-as if one category of mind would almost work against the rationality of contingency.

How to explain this within the realms of science?

One after-beer explanation goes as follows: that science in the last century has accurately pushed out of the door the idea of God as the matrix for the phenomenology of the world-including the origin of life. What all these "faith-in-life-movements" do is to quietly re-introduce God from the backyard window.

This picture is suggestive, but it does not really explain much. One step further is to say that this re-introduction of divine creation takes place unconsciously. What comes to mind is the notion of archetypes of the collective unconscious of C.G. Jung. An archetype is the part of the mental structure that is common to all men and that, according to Jung and his scholars (von Franz 1988; Meier 1992) represents the unconscious creative matrix of human affective and mental behaviour.

In their letters exchange (Meier 1992), the well-known physicist Wolfang Pauli and C.G. Jung discuss at length the influence of archetypical mind structures on science. In our case we would have the archetype of the sacredness of life. This would not appear with the same intensity in all men, but it would be more manifest for example in those who have or have had a Christian or generally a religious background. These collective patterns of human mind powerfully and by definition unconsciously influence our behaviour, including our mental habits.

If we would accept this picture, we would arrive at an interesting juxtaposition between rationality and irrationality in this field of science: that the power of mental archetypes naturally fights against the results of ratio—giving rise to a contradiction that is not deprived of an intrinsic social meaning.

Should one wish to become aware of the unconscious processes and to completely adhere to the world of contingency and rationality in general?

On this way, one may use and meditate on the classic citation by Monod (1971):

> "*We would like to think ourselves necessary, inevitable, ordained for all eternity. All religions, all philosophies and even part of science testify to the unwearyingly, heroic effort of mankind, desperately denying its own contingency*…"

References

Carr B (2001) Physics World, October 2001:23–25
Davies P (1999) The fifth miracle, …cited by M. Shermer in Sci, Amer. Jan. 2003, page 23. see also J.D.Barrow, F.J.Tripler (1988) Nature 331:31–34
Dyson F (1979) Disturbing the Universe, cited by M. Shermer, as above
de Duve C (1991) Blueprints for a cell. Neil Patterson Publ., North Carolina
de Duve C (2002) Life evolving: Molecules, Mind and Meaning. Oxford Univ. Press
Gould SJ (1989) Wonderful Life. Penguin Books, UK
Hawking S, Penrose R (1996) The Nature of Space and Time. Princeton University Press, Princeton NJ, USA, pp. 89–90
Hoyle F, Wickramasinghe C (1999) Astronomical origins of life—Steps toward panspermia —Preface Astrophys Space Sci 268 (1–3): VII–VIII
Jacob F (1982) The possible and the actual. Univ. of Washington Press, Seattle
Mastrapa RME, Glanzberg H, Head JN et al. (2001) Earth Planet Sc Lett 189 (1–2): 1–8 Jun 30
Meier CA (1992) Wolfang Pauli und C.G. Jung. Ein Briefwechsel, Springer-Verlag
Monod J (1971) Chance and Necessity. Knopf
Morowitz HJ (1992) Beginnings of Cellular Life. Yale Univ. Press
Szathmary E (2002) Nature 419:779
Tarter J (2001) Ann Ny Acad Sci 950: 269–275

von Franz M-L (1988) Psyche und Materie. Daimon Verlag

E-mail: luisi@mat.ethz.ch

Determinism vs. Contingency: A False Dichotomy

Sandra D. Mitchell

Department of History and Philosophy of Science, University of Pittsburgh, USA

Keywords Determinism • Origin of Life • Contingency • Necessity

In the Origins of Life literature a question is posed as to whether or not life arose on our planet by chance, or was it rather expected based on the laws governing our universe and the conditions that occurred in the deep past. For example, DeDuve posits two options: chance or necessity. Those that take the "chance" option defend a "gospel of contingency". Monod appeals to the "unfeeling immensity out which we arose only by chance" with what has populated the history of life as "frozen accidents" and Gould explores the different futures that would have occurred by "replaying life's tape". The sources of "chance" in the history of life itself and its diversity are mutation and the environments that have shaped adaptive evolution.

DeDuve suggests in contrast that "Most biologists, today, tend to see life and mind as cosmic imperatives, written into the very fabric of the universe, rather than as extraordinarily improbable products of chance." By necessity he means that given sufficient opportunity, all possibilities will eventually be realized, and as life was one such possibility, then it should be expected.

It is my view that the dichotomy of chance vs. necessity that has shaped the discussion of the evolution of life is a false one. Indeed, neither chance nor necessity is appropriately attributed to the origin of life. This argument requires a clarification of what is meant by such concepts, which makes space for a third option. Determinism is a view with a long history, finding a clear presentation in LaPlace's view that the world is so ordered that if, given a specified way things are at a time *t*, the way things go thereafter is fixed as a matter of natural law. Determinism requires a world that (a) has a well-defined state or description, at any given time, and (b) laws of nature that are true at all places and times (Hoefer 2008). If these are true than logically the LaPlacean "intelligence" would be able to predict all future states and retrodict all past states of the world—including the origin of life. However, there is little in history of science or contemporary science to defend these assumptions. The incompleteness of representation makes the requirement of a complete description of any state illusory. In addition, I have argued previously (Mitchell 2000, 2003, 2009) that the assumption of universal true laws is not satisfied by the knowledge science has discovered about our world and can blind us to the diversity of types of causal dependence that are found in nature. The other extreme, chance, is not appropriate either if by chance we mean absolute randomness. Those who have been allocated to this half of the dichotomy generally do not mean this, but rather something less extreme, such as randomness relative to the direction of adaptive evolution, for example, or unpredictability due to the unknown or vast number of contributing causes. These reflect types of causal dependence, not chance, just as the more determinate, predictable and law-like causal relations reflect more stable, or stronger forms of causal dependence. I maintain that looking at the character of the causal dependence that describes the way the origin of life stands to various physical, chemical, environmental, more ultimate and more proximate conditions better represents both what happened in the past and our understanding of it. Some causal dependencies are relatively context insensitive (whatever else is going on, the conservation of mass-energy law will hold) and some are more ephemeral,

providing constraints, rather than obligate outcomes (given the conditions of selection on a population at a given time, a range of functionally equivalent alternatives are possible).

My view is that a better conceptual framework in which to investigate the relative expectedness of the origin of life is one that embraces degrees of contingency, stability and strength among causes and their effects, and sees the unfolding of the universe in terms of nested constraints on possible futures, rather than the necessary, ineluctable coming into being of what was obligatory since the big bang. The history of chemical elements is an example. It took 1–3 million years after the big bang for temperatures to be cool enough for the simplest molecule to form. Several tens of millions of years later for star formation, where, it is believed, heavier elements were synthesized and released. Several billion years would be required for the formation of planets/solar systems and for the kind of chemistry required for life processes (Tuckerman 2006). The structure and behavior of physical matter made chemical objects and their behavior possible. But at various times in the history of the universe different chemical objects were realized. Indeed, as Seaborg argued, some physically possible chemical elements have not yet been realized, i.e. the super heavy actinides (see Seaborg 1994).

Is the origin of life necessary or by chance? My answer is neither. It was dependent on a particular trajectory of events, some parts of which were unchanged since the big bang, and others that are constituted by some realized possibilities, not necessities that were set up at the origin of the universe. How the dependencies are characterized itself depends on where in the unwinding of the history of the universe we stop to take stock and look forward. From the perspective of the big bang, life looks chancier; from the view of just before it occurred, life looks entirely expected.

References

DeDuve (2005) Singularities: Landmarks on the Pathways of Life. Cambridge: Cambridge University Press

Hoefer C (2008) "Causal Determinism" Stanford Encyclopedia of Philosophy. Edward Zalta N (ed) URL = http://plato.stanford.edu/entries/determinism-causal/

Mitchell SD (2000) Dimensions of scientific law. Philosophy of Science 67:242–265

Mitchell SD (2003) Biological complexity and integrative pluralism. Cambridge: Cambridge University Press

Mitchell SD (2009) Unsimple Truths: Science, Complexity and Policy. Chicago: University of Chicago Press.

Seaborg GT (1994) Modern Alchemy: Selected Papers of Glenn T. Seaborg. World Scientific

Tuckerman M (2006) "When Did Chemical Processes Start to Occur" Lecture 3: URL = http://www.nyu.edu/classes/tuckerman/honors.chem/lectures/lecture_1/node1.html

E-mail: smitchel@pitt.edu

Was the Emergence of Life on Earth a Likely Outcome of Chemical Evolution?

Andrew Pohorille

Exobiology Branch, NASA Ames Research Center, USA

Keywords Prebiotic Chemistry • Molecular Self-organization • Hydrophobic Effects • Protobiological Metabolism

Was the emergence of life a predictable outcome of chemical evolution on earth? Could evolution produce life very different from ours? These are one of the oldest questions in the field of the origin of life that not only have broad philosophical implications but also impact how we approach the problem from the methodological standpoint.

At this meeting, the issue was posed in terms of the dichotomy between contingency and determinism. This is not a fortunate framing because these two terms in their conventional meaning are neither mutually exclusive nor jointly exhaustive. Determinism, represented in natural sciences by Newtonian physics, relies on the assumption that every event is causally determined by a chain of previous events. In the context of the origin of life it means that once the initial conditions on the early earth have been specified further evolution follows inevitably. Considering uncertainties about conditions on the prebiotic earth, many plausible sets of initial conditions can be defined, each followed by a separate deterministic trajectory. This conventional understanding of determinism does not admit contingency. Further, it has no implications for evaluating how many sets of initial conditions lead to the emergence of life.

It appears that a better framing of the problem is as follows: given plausible sets of initial conditions on the early earth how probable and broadly spread are evolutionary trajectories that lead to life? Instead of undertaking an impossible task of specifying microscopic initial conditions for all components of the system one uses a reduced representation of this system and specify only a small set of essential macroscopic parameters, values (or ranges of values) of which can be identified, inferred or estimated from experiment, theory or historical record. The following evolutionary trajectories are still governed by laws of physics and chemistry but become probabilistic and "contingency" is admitted as variations in other variables in the system. A similar reasoning is common in other fields of science, for example in statistical mechanics. Some trajectories lead to life, perhaps in different forms, whereas others do not. Of our true interest is the ratio of these two outcomes. The issue of determinism does not directly enter the picture.

The debate about the likelihood of the emergence of life is quite old. One view holds that the origin of life is an event governed by chance, and the result of so many random events (contingencies) is unpredictable. This view was eloquently expressed by Monod (1971). In his book "Chance or Necessity" he argued that life was a product of "nature's roulette." In an alternative view, expressed in particular by deDuve (1995) and Morowitz (1992, Smith and Morowitz 2004), the origin of life is considered a highly probable or even inevitable event (although its details need not be determined in every respect). Only in this sense the origin of life can be considered a "deterministic event".

From the methodological point of view Monod's position renders research on the origins of life useless. As it is hopeless to predict numbers in roulette, it would be fool's errand to try to understand "nature's roulette". This, however, does not mean that if stochastic events played a role in the transitions from inanimate to animate matter, even an important one, their outcomes would be impossible to predict. The central requirement is that there are sufficiently strong, underlying constraints coming from physics and chemistry that act on the system. Unraveling these constraints and their effects on the origins of life is the key to the problem here. In addition, since the laws of physics and principles of chemistry are the same everywhere in the Universe, this reasoning implies some level of universality of life, no matter where it originated.

One set of constraints comes from prebiotic organic chemistry. Contrary to a "random chemistry" assumption adopted in a number of models of chemical evolution (Kauffman 1986; Dyson 1999), primordial chemistry was quite constrained. In particular, Weber (2002, 2004) has demonstrated that synthetic potential of carbon chemistry under mild aqueous conditions in the absence of catalysts is quite limited as a result of thermodynamic and

kinetic constraints. Even if the synthetic limitations of early metabolism are relaxed by introducing high-energy compounds that can capture some of the free energy released in downhill reactions and subsequently use it to drive uphill reactions, the diversity of chemical transformations remains restricted by the requirements that only −10 kcal/mol of carbon is available for biosynthesis, and that energy-rich molecules can be synthesized only by irreversible reactions with large, favorable free energies. The repertoire of possible prebiotic reactions can be further expanded to transformations that are kinetically forbidden without assistance, once enzymatic or non-enzymatic catalysis is included. Even then, many synthetic constraints remain if one makes a biochemically plausible assumption that chemical reactions in which the substrates and the products are separated by high energy barriers are less likely than reactions involving low energy barriers. A somewhat related reasoning about synthetic constraints led Smith and Morowitz (2004) to a conclusion that the reversed citric acid cycle was at the origin of metabolism.

Equally important are principles of and constraints on molecular self-organization. To self-reproduce and evolve, organic matter must self-organize into functional structures capable of responding to environmental changes. This process is based on physical rather than chemical interactions, *i.e.* interactions that do not involve the formation of chemical bonds. Folded proteins, membranes forming cell walls and the DNA double helix are examples of structures stabilized by such interactions. In every biological process these interactions are often formed and broken in response to internal and external stimuli. This requires that their strength must be properly tuned. If they were too weak, the system would exhibit uncontrolled response to natural fluctuations of physical and chemical parameters. If they were too strong biological processes would be too slow and energetically costly. Furthermore, strength of physical interactions depends critically on the solvent. Polar molecules can be dissolved in polar solvents, such as water, but not in non-polar ones. Electrostatic interactions between these molecules are reduced in polar liquids, compared to those in the gas phase, such that they become compatible with other physical interactions. In addition, water exhibits a remarkable trait that it also promotes *hydrophobic interactions* between non-polar molecules. The hydrophobic effect is responsible for self-organization of nanoscopic structures such as micelles, membranes and globular proteins. Thus, water is an excellent solvent for life mainly because it promotes self-organization of matter into structures that are sufficiently versatile, robust yet flexible to support functions of a living system. Only very few other solvents might have similar properties. This illustrates how physical interactions greatly limit environments that are suitable for life.

The constraints of physics and chemistry also act at the level of self-organization of chemical reactions into networks. We recently explored this issue through modeling the emergence of metabolism and enzymatic catalysis using chemically and biochemically plausible assumptions (Pohorille 2008). Our computer simulations indicate that although most “protocells” exhibit little catalytic activity, some encapsulate metabolisms composed of series of consecutive chemical reactions, which occasionally organize into autocatalytic cycles, even without genome. Even though the underlying processes are highly stochastic and the mathematical formulation of the model is fully probabilistic, these features are conserved in the populations. In addition, several concepts inherent to Darwinian evolution, such as the “species” (defined as similar metabolic networks), fitness to the environment and inheritance appear to hold for the population, but not for individual protocells. Lancet et al. reached a similar conclusion (Segrè et al. 2001). They showed that compositions of protocells without genomes could persist over generations, thus forming “compositional genomes”.

In summary, even though we are still far from definitive answers to the question that was posed here, a number of experimental and theoretical tools are now available that allows us to shift the problem from the realm of religion and speculations to the realm of rational, scientific inquiry.

References

De Duve C (1995) Vital Dust: Life As a Cosmic Imperative. Basic Books, New York

Dyson FJ (1999) Origins of Life. Cambridge University Press, Cambridge

Kauffman SA (1986) Autocatalytic sets of proteins. J. Theor. Biol 119: 1–24

Monod J (1971) Chance and Necessity. Alfred A. Knopf, New York

Morowitz HJ (1992) Beginnings of Cellular Life. Yale University Press, New Haven

Pohorille A (2008) 'Protocells as Universal Ancestors of Living Systems', in S. Rasmussen, M. Bedau, L. Chen, D. Deamer, D. Krakauer, N. Packard, and P. Stadler, (eds) Protocells: Bridging Nonliving and Living Matter, MIT Press, Cambridge, MA., pp 563–582

Segre D, Shenhav B, Kafri R, Lancet D (2001) The molecular roots of compositional inheritance. J Theor Biol 213: 481–491

Smith E, Morowitz H J (2004) Universality in intermediary metabolism. Proc Natl Acad Sci USA 101: 13168–13173

Weber AL (2002) Chemical constraints governing the origin of metabolism: the thermodynamic landscape of carbon group transformations under mild aqueous conditions. Orig Life Evol Biosphere 32: 333–357

Weber AL (2004) Kinetics of organic transformations under mild aqueous conditions: implications for the origin of life and its metabolism. Orig Life Evol Biosphere 34: 473–495

E-mail: Andrew.Pohorille@nasa.gov

On Question 1: Contingency vs. Determinism

Víctor Norris, Anthony Delaune

AMMIS Laboratory EA 3829, University of Rouen, France

Keywords Molecular Complementarity • Composome • Prebiotic Ecology • Hyperstructure

Introduction

The molecular biologist seeking to understand phenotypes is confronted with an intracellular interaction network comprising thousands or tens of thousands of different genes and hundreds of thousands, even a million, different proteins in their various modified forms. It is hardly surprising therefore that molecular biologists tend to believe that the origin of the cell was largely the result of chance events (Monod 1971). The problem with this view is that it is a cop-out. It offers no real explanation and closes the door on the search for other explanations. In contrast, the determinist view (de Duve 1995), far from being 'crypto-creationist', has the merit of opening this door and of encouraging

reflection and experimentation. Here, we espouse determinism and reject the notion that life arose as a miraculous, rare, chance event. Instead, we propose that the problems that living systems were (and still are) obliged to solve constrained the space of possibilities open to life. These constraints on possibilities were so severe that the deterministic elements responsible for the origins of life can be identified.

Life's problems

There are three related problems that life must solve. The first entails choosing between competition and collaboration. The difficulty in making this choice is that there are many circumstances in which it is better to choose competition and to succeed at the expense of other cells yet there are many other circumstances in which it is better to choose collaboration and to make concessions for the common good. The second problem entails choosing between growing in heaven or surviving in hell. Cells that can outgrow others and produce more progeny should replace their rivals. But growth makes cells vulnerable to environmental disasters—and a dead cell has no progeny. An alternative strategy is to stop growing and become, for example, a spore, but again a non-growing bacterium has no progeny. The third problem entails choosing between specialising and diversifying. There are circumstances in which it is better to concentrate resources on exploiting a single niche with great efficiency and other circumstances in which it is better to distribute resources so as to exploit many niches with lesser efficiency.

Bacteria have solved these problem by operating as a *population* of interacting cells in which the different phenotypes of the cells are generated by progression through the cell cycle. Collaboration is achieved whereby competent cells can undergo autolysis to release DNA that benefits other cells of the same species (Vollmer et al. 2008). Choosing between growth and survival strategies is achieved because DNA has two strands, one of which in bacteria such as *Bacillus subtilis* tends to contain the genes needed for growth whilst the other strand contains the genes needed for survival (Rocha et al. 2003). This means that during the cell cycle, the replication of DNA and subsequent DNA segregation and cell division can generate progeny with complementary phenotypes (Norris et al. 2002; Norris and Fishov 2001; Norris and Madsen 1995). These complementary phenotypes precondition the population so that it contains a diversity of specialists. Hence, some cells in a population can spontaneously stop growing and escape the unexpected addition of an antibiotic that destroys other genetically identical cells in the same population (Balaban, et al. 2004); the non-growing cells can later resume growth in the absence of the antibiotic to regenerate the population.

The first constraint—molecular complementarity

In the 'ecosystems first' scenario, we have proposed that life evolved as a diverse interacting community of molecules from the start, and not as a single replicating entity (Hunding et al. 2006; Norris et al. 2007; Root-Bernstein R.S. 2009). More specifically, Life originated as an ecosystem of composomes exchanging constituents in an abiotic flow of creation and degradation. These composomes, to adopt a term coined by Lancet and collaborators (Segre et al. 2000), were non-covalent assemblies of molecules that interacted under the constraints of molecular complementarity, to adopt the term coined by Root-Bernstein and collaborators (Root-Bernstein R. S. and Dillon 1997) and that evolved via fission-fusion (Norris and Raine 1998) and via catalysis of the production of their constituents (Segre et al. 2000). There is an echo here of the 'islands' of Dyson (Dyson 1982) but note that our composomes combined with one another. Moreover, and this is important for our argument here, composomes were initially formed and maintained by

interactions between constituents based on molecular complementarity that protected abiotically created molecules from degradation. Hence, composomes initially grew by preserving molecules in a flux of abiotic creation and degradation, thereby allowing these molecules to accumulate as compositionally distinct composomes (or structurally distinct pre-composomes) *without the need for catalysis* (Norris and Raine 1998). In other words, structure came first! In this way, molecular complementarity constituted an enormous constraint on the origins of life, helping to take these origins out of the realm of contingency and into that of determinism.

The second constraint—frequent catalysis

A second major constraint on the origins of life in the ecosystems-first scenario was that when catalysis did start to become to be important it was frequent and non-specific rather than rare and specific. For example, segregation of lipids into different domains on the surface of the composome would have created an interface or boundary between these domains. Such an interface would have an affinity for amino acids and by concentrating, aligning and orienting them, favour peptide bond formation (Raine and Norris 2007). A similar argument can be made for nucleotide bond formation. Thus the dynamic and varied interfaces between domains on composomal surfaces acted as the first ribosomes and polymerases. This catalytic activity was relatively non-specific and relatively inefficient. Other composome constituents, such as polyphosphate and polyhydroxybutrate (Norris 2005), also formed phases with catalytic activities that were again relatively non-specific but frequent. In this way, the dynamic and varied structures formed by composome constituents catalysed production of the myriad constituents of the composome or of other composomes in the population. The composomal population became increasingly enriched in molecules that interacted with one another to form catalytic structures that catalysed a large number of reactions. The argument here is in favour of determinism because initially a huge range of species of composomal molecules participated in relatively non-specific catalyses and this range was progressively limited as catalysis became increasingly specific.

The third constraint—from complicated to simple

A third major constraint in the eco-systems first scenario was that of producing the simple from the complex and complicated. How did life get reproducible, selectable behaviours out of many, varied constituents? This problem has been visited by Kauffman in terms of (for example) the 4,000 genes of *Escherichia coli* generating a phenotype space of 2^{4000} on-off combinations of those genes (Kauffman 1996). The constraint is one of *coherence*—coherence at both the present and successive times. Different enzymes are produced that function together in a way that is coherent with respect to one another and to the environment—to give, in other words—a *meaningful* phenotype. This phenotype is preserved over time so that there is a continuity of behaviours—to give, in other words—a meaningful constancy. We have argued in the fission-fusion model that one of the key events of the cell cycle, division, is a key to producing distinct, complementary, coherent phenotypes (Norris and Raine 1998). Coherence can only be achieved by interactions hence the necessity for life to be coherent imposed a powerful constraint on the nature and variety of molecules and on their interactions in composomes.

Relevance to modern cells

The determinist view of the origins of life is firmly rooted in the reality of modern cells and populations. After decades of controversy (for references see (Norris, et al. 1994)),

it is now clear that both prokaryotic and eukaryotic cells are stuffed with structures (Cheng et al. 2008; Narayanaswamy et al. 2009; Norris et al. 2007; Watt et al. 2007). Many of these structures can be considered as hyperstructures where a hyperstructure is defined as a large assembly of many, interacting, often unstable molecules and macromolecules that performs a function. Interactions between enzymic hyperstructures creates an enzoskeleton (Norris et al. 1996) and *functioning-dependent* interactions between enzymes that depend on the presence of their substrates have important implications for signalling (Janniere et al. 2007; Thellier et al. 2006). Yet, although there are few macromolecules that are not in structures in modern cells, contingent approaches are founded on unstructured cells that do not exist—and probably never did.

It could be argued that much of molecular biology is concerned with the study of a contingent veneer that overlays deeper, deterministic principles. Consider for example the number of studies devoted to the complex roles of the DnaA protein and other proteins in the control of initiation at the chromosomal origin of replication in *E. coli*, *oriC*. This initiation depends the local opening of the duplex, a key step, which involves DnaA binding to DnaA boxes clustered within *oriC*. It turns out, however, that initiation of replication can occur in certain mutants that is independent of both DnaA and *oriC* (Kogoma 1997). This initiation requires a mutation in an RNase such that transcription lasts long enough to form an R-loop and displace the other DNA strand. At a deep level, in other words, transcription is coupled to initiation of replication via the physical event of duplex opening.

The determinist view is based on the collective properties of modern cellular constituents which have explanations in terms of architecture and physical chemistry rather than the simple molecular biology which underpins contingency (Hunding et al. 2006). The essentially integrative approach of a prebiotic ecology greatly narrowed the range of possibilities, thereby shifting life's origins into the realm of determinism.

The determinist view that life has to be *coherent* imposes a powerful constraint on the nature and variety of molecules and on their interactions in modern cells (Norris et al. 2005). We have argued that the modern cell cycle generates populations of bacterial cells with coherent phenotypes. In the strand segregation hypothesis (Rocha et al. 2003), the separate sets of hyperstructures associated with the two strands of DNA allow DNA replication, segregation and cell division to produce daughter cells with complementary, coherent phenotypes; recent evidence is consistent with this hypothesis (White, et al. 2008). The necessity for life to be coherent plays a similarly important role in the prebiotic ecology scenario (Hunding et al. 2006; Norris et al. 2007).

Discussion

The principal difference between the deterministic scenario of a prebiotic ecology and the contingency hypothesis is that the former offers *explanations*. The importance of the composome in life's early solutions is reflected in the importance of the hyperstructure in life's recent solutions (Norris et al. 2007). The importance of fission-fusion in the evolution of composomes is reflected in the importance of the cell cycle in modern cell populations (Norris et al. 2004; Norris et al. 2002; Rocha et al. 2003). Molecular complementarity was as important in life's origins as it is now (Root-Bernstein and Dillon 1997).

Rather than being an arbitrary, one-off event, life is, in our determinist scenario, inevitable and predictable. Rather than being alone on Earth, life abounds elsewhere in the Universe. Rather than it being impossible to rerun life's origins, recreation becomes feasible. Rather than an absence of explanation and an allegiance to a god of 'blind chance',

determinism offers common themes and deep answers that ultimately will form part of a unifying theory of biology. Insofar as contingency did play a role in the origins of life, it was in only in the initial selection of exactly which molecules interacted. The direction that life took as it evolved is largely determined by the composomal compass even if minor vagaries occurred that were contingent on the local terrain.

In summary, we would argue that choosing between contingency and determinism ought to depend on answering the question 'what is a cell' (Norris et al. 2004)? And what we see now is that a modern bacterial cell is a set of hyperstructures and that this cell is itself part of a population (Norris et al. 2007).

References

Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S (2004) Bacterial persistence as a phenotypic switch. Science 305:1622–5

Cheng S, Liu Y, Crowley CS, Yeates TO, Bobik TA (2008) Bacterial microcompartments: their properties and paradoxes. Bioessays 30:1084–95

de Duve C (1995) Vital dust: Life as a cosmic imperative. New York: Basic Books

Dyson FJ (1982) A model for the origin of life. J Mol Evol 18:344–50

Hunding A, Kepes F, Lancet D, Minsky A, Norris V, et al (2006) Compositional complementarity and prebiotic ecology in the origin of life. Bioessays 28: 399–412

Janniere L, Canceill D, Suski C, Kanga S, Dalmais B, et al (2007) Genetic evidence for a link between glycolysis and DNA replication. PLoS ONE 2:e447

Kauffman S (1996) At home in the Universe, the search for the laws of complexity. London: Penguin. 1–321 pp.

Kogoma T (1997) Stable DNA replication: interplay between DNA replication, homologous recombination and transcription. Microbiology and Molecular Biology Reviews 61:212–38

Monod J (1971) Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology. New York: Alfred A. Knopf

Narayanaswamy R, Levy M, Tsechansky M, Stovall GM, O'Connell JD, et al. 2009. Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation. Proceedings of the National Academy of Sciences of the United States of America 106:10147–52

Norris V (2005) Poly-(R)-3-hydroxybutyrate and the pioneering work of Rosetta Natoli Reusch. Cellular and molecular biology (Noisy-le-Grand, France) 51:629–34

Norris V, Amar P, Bernot G, Delaune A, Derue C, et al (2004) Questions for cell cyclists. Journal of Biological Physics and Chemistry 4:124–30

Norris V, Ayala JA, Begg K, Bouché J-P, Bouloc P, et al (1994) Cell cycle control: prokaryotic solutions to eukaryotic problems? Journal of theoretical biology 168:227–30

Norris V, Blaauwen TD, Doi RH, Harshey RM, Janniere L, et al (2007) Toward a Hyperstructure Taxonomy. Annual review of microbiology 61:309–29

Norris V, Cabin A, Zemirline A (2005) Hypercomplexity. Acta biotheoretica 53:313–30

Norris V, Demarty M, Raine D, Cabin-Flaman A, Le Sceller L 2002. Hypothesis: hyperstructures regulate initiation in Escherichia coli and other bacteria. Biochimie 84:341–7

Norris V, den Blaauwen T, Cabin-Flaman A, Doi RH, Harshey R, et al (2007) Functional taxonomy of bacterial hyperstructures. Microbiology and Molecular Biology Reviews 71:230–53

Norris V, Fishov I (2001) Hypothesis: membrane domains and hyperstructures control bacterial division. Biochimie 83:91–8

Norris V, Hunding A, Kepes F, Lancet D, Minsky A, et al (2007) Question 7: the first units of life were not simple cells. Orig Life Evol Biosph 37:429–32

Norris V, Madsen MS (1995) Autocatalytic gene expression occurs via transertion and membrane domain formation and underlies differentiation in bacteria: a model. Journal of molecular biology 253:739–48

Norris V, Raine DJ (1998) A fission-fusion origin for life. Origins of Life and Evolution of the Biosphere 28:523–37

Norris V, Turnock G, Sigee D (1996) The Escherichia coli enzoskeleton. Molecular microbiology 19:197–204

Raine DJ, Norris V (2007) Lipid domain boundaries as prebiotic catalysts of peptide bond formation. Journal of theoretical biology 246:176–85

Rocha E, Fralick J, Vediyappan G, Danchin A, Norris V (2003) A strand-specific model for chromosome segregation in bacteria. Molecular microbiology 49:895–903

Root-Bernstein RS (2009) An "ecosystems-first" theory of the origin of life based on molecular complementarity. In Astrobiology: Emergence, Search, and Detection of Life, ed. VA Basiuk, pp. 1–30: American Scientific Publishers

Root-Bernstein RS, Dillon PF (1997) Molecular complementarity I: the complementarity theory of the origin and evolution of life. Journal of theoretical biology 188:447–79

Segre D, Ben-Eli D, Lancet D (2000) Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. Proceedings of the National Academy of Science USA 97:4112–7

Thellier M, Legent G, Amar P, Norris V, Ripoll C (2006) Steady-state kinetic behaviour of functioning-dependent structures. The FEBS journal 273:4287–99

Vollmer W, Joris B, Charlier P, Foster S (2008) Bacterial peptidoglycan (murein) hydrolases. FEMS microbiology reviews 32:259–86

Watt RM, Wang J, Leong M, Kung HF, Cheah KS, et al (2007) Visualizing the proteome of Escherichia coli: an efficient and versatile method for labeling chromosomal coding DNA sequences (CDSs) with fluorescent protein genes. Nucleic acids research 35:e37

White MA, Eykelenboom JK, Lopez-Vernaza MA, Wilson E, Leach DR (2008) Non-random segregation of sister chromosomes in Escherichia coli. Nature 455:1248–50

E-mail: victor.norris@wanadoo.fr

An Epistemology of Contingency: Chance and Determinism at the Origin of Life

Leonardo Bich, Gianluca Bocchi, Luisa Damiano
Post Graduate School of Anthropology and Epistemology of Complexity, University of Bergamo Italy; Dipartimento di Scienze della Persona and Post Graduate School of Anthropology and Epistemology of Complexity, University of Bergamo Italy; Graduate School of Core Ethics and Frontier Sciences, Ritsumeikan University, Japan

Keywords Epistemological Constructivism • Practical Contingency • Strong Contingency • Trans-computability • Non-computability

1. Introduction

By the adjective contingent, we characterize those phenomena which cannot be included in deterministic causal chains. This meaning of contingency is frequently found in the contemporary scientific thought. Increasingly often it is used in the definition of the mechanisms which in nature generate novelties—be they positive or disruptive ones—or variety (Bocchi and Ceruti 1993, 2009; Oyama et al 2001). Among the well-known examples of this use we can include the "catastrophic" hypothesis of the extinction of dinosaurs due to a collision of an asteroid with the Earth, and the Monodian explanation of variability as the result of random mutations (Monod 1970).

In the following pages we answer the question of "contingency versus determinism" through a brief epistemological analysis of the different applications of the concept of contingency in the scientific thought. The result consists in the rejection of the statement according to which it is impossible to decide in a great deal of cases whether to the phenomenon under scrutiny should be applied the label of "contingent" or "deterministic" on a scientific basis. More precisely our analysis leads us to propose the thesis that this choice can be done on a heuristic basis, that is, through a criterion of descriptive pertinence. What looks like a hypothesis about Nature herself from the more traditional objectivist points of view, from the perspective we adhere to and we'll try to discuss in the following remarks becomes a hypothesis on the limits of our attempts to describe her (Prigogine and Stengers 1979; Maturana 1988; Bich 2008; Damiano 2009).

This proposal contains a methodological suggestion for contemporary science: to consider determinism and contingency not as "ontological properties" of natural world, referred to reality itself, but rather as "heuristic properties", which an observer ascribes to nature in order to provide a description able to allow efficacious interactions with it. This methodological principle is what we propose to the studies on the origin of life, in order to emancipate them from the needs of belief in the choice between determinism and contingency. According to this perspective the problem of this choice consists in trying to establish what model is more pertinent, that is, as we will show, not objectively true but descriptively efficacious. On this point the answer is provided not only from the analysis of models, but also and especially by the interactions they allow with reality: empirical experiences which today are also made possible by synthetic biology and artificial life. The answer we propose to the question is to focus on the empirical corroborations of the models of contingency and determinism, in order that the choice would not be arbitrary but consistent with the procedure of experimental knowledge.

2. Three versions of contingency

The issue "contingency versus determinism" has received some attention in various fields of physical research and also in the computational domain in its relation with randomness. On the contrary, its analysis in the biological domain should be conducted with a more in-depth attitude (Longo 2009). Above all, this situation is related to the extreme complexity of the biological domain, characterized by the intertwined interaction of different levels of organization, and also to the lack of detailed theoretical and formal models of living systems.

If we attempt to put some order in the issue of contingency, we can identify at least three different applications of this concept in scientific descriptions. All of them have a peculiar characteristic which can be summarized as follows: the contingency to which

they refer is always susceptible of being considered not as an objective property of reality, but rather as the projection on reality of a subjective aspect: the incapability of the observer to provide a deterministic description. When we're considering these uses, we can ask which kinds of limits of our descriptive models they can entail and, furthermore, whether there is any effective possibility to bring them back to a deterministic framework and if there are any, which is the meaning of this sort of reduction for our idea of knowledge.

The first use of the concept of contingency applies to those events whose origin can be considered as wholly unrelated to the object of our modelization and thus accidental from the point of view of a causal description: they are not included in the original extension of a model of the system considered. Surely, the previous example of a celestial phenomenon deeply influencing the biotic and climatic ones (the collision with an asteroid) as well as the hypotheses of any celestial trigger for the origins of life on Earth, can be considered as belonging to this first class of contingent phenomena.

It is easy to see how the "contingency" hypothesized in this kind of scientific explanations can be always considered, rather than a characteristic of the phenomena under investigation, as the expression of the incapability of the observer to define a comprehensive deterministic explanatory mechanism. A transition from a more contingency-oriented to a more determinism-oriented explanation can be envisaged, at least in principle. In these cases there is no theoretical limit to our possibility to build a more comprehensive description by increasing the resolution power of the starting model, so to incorporate the explanation of the problematic event. We can conceive even an ongoing process with several steps, each of them corresponding to a more comprehensive deterministic model: a potentially infinite process of reduction of the ignorance of the observer.

The second use of contingency refers to the very wide class of processes whose temporal evolution is inherently unpredictable due to the high number of factors implicated and, above all, to the complexity of the dynamical patterns regulating the behavior of such systems. If we start with a range of possible pathways that the system can choose in its evolution, it is impossible to say in advance which one will be chosen. Furthermore, if we to look backwards at the temporal evolution of a systems, it is not possible to explain in a great deal of cases why a specific pathway and not another was chosen. This type of contingency is indeed a very widespread phenomenon in living systems: we can only think of the so-called "frozen accidents", crucial threshold of evolution in which a peculiar pathway, that has originally developed in an unpredictable way and without apparent selective advantage with respect to other possible pathways, becomes fixed and exclusive, thus constituting the initial conditions for further evolution. A wide class of this second type of contingency is constituted by the so called *ex-aptations* (Kauffman 2008): structural or functional developments of biological systems that are not the best solutions that can be envisaged on the basis of natural selection alone, but that results from a very complex balance between heterogeneous and often opposite needs and factors.

For this class of phenomena we can speak of a "practical" contingency. Although the dynamics which characterize the behavior of this kind of systems in our models are usually deterministic, such as in the very popular models of deterministic chaos, and we can say that the various possible results are in principle pre-determined, these models do not provide a specific evolutive solution (a single trajectory), but only a probabilistic one (multiple trajectories). Here chance is not dispensable because of our intrinsic limitations as observers in measuring the initial conditions and in considering all the elements necessary in order to describe the process in its entirety. They are "practical limits" because they concern our

observational capabilities; but they are limits "in principle" because cannot be overcome by any improvements of our observational tools: with Bedau we can say that they involve a practical limitation in principle (Bedau 2008).

We can therefore identify there a sort of convergence between determinism and contingency as deterministic models sustain contingentist descriptions. In fact, it is problematic to consider contingency, as well as determinism, as properties of reality. The fact that deterministic tools give rise to contingentist outcomes seems to remove the possibility to clearly establish if reality is contingent or deterministic. This induces us to think that it is neither one nor the other. Simply, in order to describe it we need to combine deterministic and contingentist attitudes.

This conclusion is strengthened by the analysis of the third use of the notion of contingency. It is a stronger one, which can be associated to what Monod calls "gratuity" (Monod 1970). It entails a limit in principle in any possible model we can trace of a phenomenon, when it cannot be derived from some given initial conditions. This depends on the fact that both the dynamics and the range of evolutionary possibilities cannot be pre-established. At different evolutive steps new emergent elements and dynamics involving them are observed, in such a way that the whole process cannot be considered a simple exploration of a given space of evolutionary pathways and possibilities (Kampis 1991; Kauffman 2008; Longo 2009). For example if we consider oxygen as a powerful factor in promoting multicellularity and animal diversity, above all in the era of the so-called "Cambrian explosion" (Bocchi and Ceruti 2009), we should also bear in mind that oxygen was a very toxic gas for the microorganisms that populated the Earth in the very first eons of her history, in the first half of the three-billions-eight-hundred-millions long history of life. The transition from a biosphere without oxygen to a biosphere with oxygen as a powerful motor of change and diversification surely was wholly unpredictable and underivable from any condition of the ancient eons of evolution. Another example, to which Monod refers with the concept of "gratuity", concerns the apparent impossibility of deriving the origins of the functions of proteins from that of their respective structures: this connection is only a matter of historical developments and does not reflect any selective advantage of a given structure with respect to the others.

In these cases the phenomena are considered as contingent with respect to their initial conditions because not derivable from them, whichever model we are willing to adopt. This radical underivability from initial conditions and dynamics expresses the absolute impossibility to produce a completely deterministic description. In this way we realize that determinism (the idea that reality is deterministic) is a not an objective truth, but a theoretical postulate founding a descriptive approach—an approach that in cases like these is not productive.

From this standpoint, the analysis of the last two applications of contingency provides also arguments against a strong "computational" view of science—characterized by a deterministic assumption and by the use of algorithmic models as the privileged tools in order to describe reality—thus opening the way to a scientific practice which promotes the use of a plurality of qualitatively different models. The natural world, and above all the history of living systems, in fact, are very rich in phenomena which, from a descriptive point of view, we can consider as "trans-computational" (Kauffman 2008): they are beyond any practical possibility of computation not only by the single human mind, but even by the most refined and powerful web of super-computers we can conceive of. In the case of the hypothesis of strong contingency, when it happens to be consistent from the theoretical and empirical points of view, we need even to describe the phenomena under scrutiny through non-algorithmic models, as none of them can be derived from any initial description (Rosen

1991; Bich 2008; Kauffman 2008). However, this latter case is still under discussion, and a result of paramount importance would be a formal demonstration of the impossibility of a derivation of some crucial step in evolution from the previous ones (may be of some crucial step leading to the life as we know it on Earth). This would constitute a crucial step towards the development of a practice of multiple models in biology, based on a "negative theorem" of non-computability may be comparable to the fundamental ones formulated in physics and mathematics during their historical developments in the last century (Bailly and Longo 2006).

3. Concluding remarks: towards a practice of contingency

The account of these three applications of the concept of contingency to major changes and transitions in evolutionary processes implies a precise diversification in the limits of our implementation of models aimed at explaining historical events and developments we testify in evolution: limits in fact but not in principle, practical limits in principle, but also theoretical limits in principle. These limits, expressing the incapability of producing deterministic descriptions and the necessity to recur also to contingentist approaches, provide a good reason to renounce to ascribe to the natural world an ontological status of contingency or determinism. If we attempted to do it, we would have to assert that reality in some aspects is deterministic, in others contingent and other both of them, depending on the descriptive tools we are using at the moment according to our purposes and points of view. Instead of considering "contingency versus determinism" as an ontological issue ("is reality itself deterministic or contingent?"), we can conceive it just as a heuristic problem ("at the moment which is the better way to describe it?"). In this way we are able to treat it as a question to which we can answer scientifically: on the basis of an evaluation of pertinence—that is, coherence and experimental effectiveness—of the models we are using.

According to these remarks our suggestion to life sciences is not to choose one of the two concepts (determinism and contingency) in the explanation of the origin of life, but instead to focus on the analysis of models and on their pertinence in this domain of research. Furthermore at the present state of biological sciences an empirical evaluation of modelizations can be achieved not only through the traditional biological investigation, but also through the experimentations provided by the science of the artificial, in particular by synthetic biology and artificial life.

References

Bailly F, Longo G (2006) Mathématiques at sciences de la nature. La singularité physique du vivant. Hermann Éditeurs, Paris

Bedau M (2008) Is Weak Emergence Just in the Mind? Minds & Machines 18:443–459

Bich L (2008) L'ordine invisibile. Ripensare il vivente oltre la metafora computazionale. Ph.D. dissertation, University of Bergamo

Bocchi G, Ceruti M (1993) Origini di storie. Feltrinelli, Milano

Bocchi G, Ceruti M (2009) Il fuoco e il gelo. Vincoli e possibilità della vita. Cortina, Milano (in press)

Damiano L (2009) Unità in dialogo. Mondadori, Milano

Kampis G (1991) Self-Modifying Systems in Biology and Cognitive Science. Pergamon Press, Oxford

Kauffman S (2008) Reinventing the Sacred. Basic Books, New York

Longo G (2009) Randomness and Determination, from Physics and Computing towards Biology. In: Nielsen M, Kucera A, Miltersen PB, Palamidessi C, Tuma P, Valencia F (eds) Lecture Notes in Computer Science, SOFSEM 2009: Theory and Practice of Computer Science, Springer, New York, pp. 49–61

Maturana H (1988) Reality: the Search for Objectivity or the Quest for a Compelling Argument. Irish J Psychol 9(1):25–85

Monod J (1970) Les hasard et la necessité. Seuil, Paris

Oyama S, Griffiths PE, Rusell DG (eds) (2001) Cycles of Contingency. Developmental Systems and Evolution. The MIT Press, Cambridge MA

Prigogine I, Stengers I (1979) La Nouvelle Alliance. Métamorphose de la science. Gallimard, Paris

Rosen R (1991) Life Itself: a Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life. Columbia University Press, New York

E-mail: leobich@libero.it

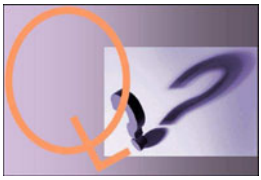
Workshop OQOL'09

Extended Abstracts for the Following Selected Question

- **Is Life an Emergent Property?**

Premise. Although emergence is a notion with many complex sides, the general view is that emergent properties are those novel properties that arise when parts or components assemble together into a higher hierarchic order—novel in the sense that they are not present in the parts or components. Most of modern scientists would consider cellular life an emergent property, as the single components are 'per se' not living. Then…

> **The question**. Do you think there are sufficient data now to say that life is indeed an emergent property, arising from the interactions and self-organization of non living parts?—Or do you still see a kind of "vitalistic" flavour in the statements that define life as an emergent quality?



Gödel, Biology and Emergent Properties

Andrés Moya

Instituto Cavanilles de Biodiversidad y Biología Evolutiva Universitat de València, Spain

Keywords Gödel's Theorems • Gödel-Turing-Chaitin (GTC) Limit • Turing Machine • Cell System • Limit to Knowledge • Emergent Properties • Biological Evolution

Gödel theorems (1931) have been quite influential in formal sciences and Philosophy. Later, Turing's halting problem (Turing 1936) and Chaitin's omega number (Chaitin 1974) have demonstrated that there are limits to what can be known formally and proven true. The Gödel-Turing-Chaitin (GTC) limit represents the threshold to what we can know regardless of the amount of time and space we are willing to invest to acquire such knowledge. Indeed, for any given formal system, there are statements that cannot be shown to be true or false or other statements that are true or false but from which we cannot derive the corresponding proof. Can we apply this GTC limit to physical (biological) entities? What are the consequences? This application is much more than a pure intellectual exercise because certain relevant behaviours and properties of material entities could be understood if we admit the existence of such limit (Ben-Jacob 1998; Danchin 2009; Penrose 1989).

Take for instance Gödel's theorems. The first Gödel theorem of undecidability applied to any biological system states that "within a cell, properties exist that are neither provable nor disprovable on the basis of the rules that define the system". It means that, on the basis of the rules governing cell behaviour, properties might exist or appear which we cannot tell whether they may or may not be derived from the rules of the cell system. The second theorem, the incompleteness theorem states that "in a sufficiently well known cell in which decidability of all properties are required, there will be contradictory properties". The biological translation of that theorem is extremely important because it asserts that, no matter how well we know a particular cell system, we can find properties and/or behaviours that seem to contradict each other. These two theorems applied to living cells, for instance, can also be interpreted within the framework of the universal Turing machines as follows: "there may appear traits or behaviours of living cells that cannot be computed by any logical machine". This statement presupposes that physical or biological entities can in principle be described in terms of an algorithm. Then, the execution of the algorithm in a finite sequence of steps, provided we have unlimited amounts of time and storage space, will allow cell behaviour computation. If Gödel theorems or the GTC limit apply to physical and/or biological entities, they tell us that we cannot anticipate the appearance of new properties in the cell or the lack of them and sometimes properties will appear that follow contradictory trajectories, no matter how deep our knowledge of the cell.

A key question in Biology is emergence. Biological systems have plenty of emergent properties and Evolutionary Biology, in particular, shows us how frequently and abundantly these occur throughout the history of living beings on Earth. The emergence of evolutionary novelties, at least some of them, gives support to Gödel's statements. Biological features, particularly those of evolutionary nature, are not predictable most of the time, and can be considered as emergent novelties/properties within that particular system formed by living entities on Earth. In summary, emergent biological phenomena may appear within a particular system that follows Gödel's theorems. We can define a new biological system adding new rules that may integrate the emergent feature, to avoid inconsistencies. However, according to Gödel's statements, the new system, although more sophisticated and far-reaching, will again come up against new unpredictable phenomena.

We can consider the evolutionary process as an executable algorithm which tells us, with precision, all the forces (rules) acting on populations of living and genetically diverse objects (agents). Let us also assume that we have, as stated before, unlimited amounts of time and storage space, further that the rules and agents are implemented in a given algorithm following current top-down (deterministic) or bottom-up (machine learning) computational approaches (Penrose 1989). Can we predict any expected outcome of such a process? The answer is "no"

according the GTC limit. By this I am not stating that life is completely unpredictable in all instances. It often is for a certain number of behaviours and traits but, from time to time, through evolutionary history, emergent phenomena have appeared. Emergent phenomena in living beings are equivalent to the unexpected outcomes of the algorithmic evolutionary process. It seems that evolution, emergent phenomena and the unpredictability of the history of life as a whole are perfectly compatible with Gödel's statements.

Let us suppose that we add new rules and agents to our first evolutionary algorithm, in such a way that we are now in a position to explain that particular phenomenon, which was an emergent phenomenon within it. Then we can obtain a new evolutionary algorithm. Although this is more sophisticated and far-reaching than the former, it will be exposed, following GTC limit, to new unpredictable phenomena.

One last remark. Contrary to interpreting Gödel's statements and GTC limit in the negative sense of a threshold to our capacity of knowledge (Moya et al. 2009), what we observe is the intrinsic ability of particular material systems, particularly living ones, to permanently create new information and then to evolve. This is possible because at a given moment in the early stages of evolution of life, a living device emerged, formed by a unit of coded information (DNA) and another device (the protein machinery) that decodes and recodes the genetic information. Within the context of research into the origin of life, the living device I am talking about is both rather complex and emergent, which may appear from a putative algorithm formed by a set of rules and agents that represent our state of knowledge on the topic. Certainly, we have not reached this state yet and research will continue to gain and appropriate knowledge of the steps and components necessary to synthesize primitive life. But there may be a second stage in such research: when life itself may be predicted on the basis of reformulating the previous algorithm, adding critical new rules and/or agents. This new algorithm, however, will not rule out the appearance of other emergent properties if we agree that it, like any other entity, is governed by the GTC limit.

Acknowledgements
The research leading to these results has received funding from the European community's Seventh Framework Programme (FP7/2007–2013) under grant agreement n° 212894, TARPOL consortia.

References

Ben-Jacob E (1998) Bacterial Wisdom, Godel's Theorem and Creative Genomic Webs. Physica A 48:57–76

Chaitin GJ (1974) Information-theoretic limitations of formal systems. JACM 21:403–424

Danchin A (2009) Bacteria as computers making computers. FEMS Microbiol Rev 33:3–26

Gödel K (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. Monatshefte für Mathematik und Physik 38:173–198

Moya A, Krasnogor N, Peretó J, Latorre A (2009) Goethe's dream. Challenges and opportunities for Synthetic Biology EMBO reports, in press.

Penrose R (1989) The emperor's new mind: concerning computers, minds, and the laws of physics. Oxford University Press, Oxford

Turing A (1936). On computable numbers, with an application to the Entscheidungs problem. Proc Lond Math Soc Series 2(42):230–265

E-mail: andres.moya@uv.es

The Chemistry that Preceded Life's Origin: When is an Evolutionary Story an Emergent Story?

Sandra Pizzarello

Arizona State University, USA

Keywords Prebiotic Evolution • Carbonaceous Meteorites • Antarctica Meteorites • Ammonia • Amines • Amino Acids

It seems that the question may be answered, at best, only abstractedly. Because there is exceedingly little knowledge about the actual origin(s) of life, we have reasonably extended to its beginnings the same fundamental evolutionary nature that has been recorded throughout life's cellular history. Albeit *a posteriori*, this reasoning warrants the general assumption that life was preceded by a prebiotic molecular evolution, where "non living parts", such as the simple molecules that multiple sources made available on the early Earth, reacted into more complex chemical structures and, of these, some attained selected associations, desirable interactions, self-copying preservation and eventually "emerged" in the self reproductive and evolving systems we call life. In these terms, the basic question in the study of the origins of life is not whether life emerged but when and how. A host of more questions easily follow about which particular molecules could have learned first such useful assembling, interacting and so on. Eventually, therefore, the proofs that "interactions and self-organization of non living parts" could be systemic and prebiotically relevant must be analytical. Analytical approaches may never get to learn the very paths that terrestrial life followed for its emergence (Eschenmoser 2007) but they alone may yet recognize the likely physico-chemical processes where "novel properties…. arise when parts or components assemble together into a higher hierarchic order".

That is, in fact, what has been pursued for close to a century as factual evidence was gathered up and down the time lines of both evolutionary biology and cosmochemistry. A still fertile field for information about chemical evolution has been the study of carbonaceous chondrites (CC). The analyses of these carbon-containing meteorites have long been part of the discourse on the origins for life, given that they are the only extraterrestrial sample available in significant amounts and their direct analyses offer a realistic view of the organic chemistry that, by either analogy of formative processes or direct delivery, could have preceded the emergence of life on the early Earth.

The prebiotic story told by meteorites started just about 40 years ago with the fall of the Murchison meteorite in 1969. We have learned since that the chemical evolution of the biogenic elements ahead of life may form abundant a-biotic organic materials as diverse as insoluble kerogen-like macromolecules and smaller compounds of the type present in biochemistry. From the deuterium and ^{13}C enrichments of Murchison's amino acids, we have also come to realize that this evolution is rooted in presolar cosmochemical environments and, from the L-enantiomeric excesses (ee) displayed by some of these same amino acids, we know now that molecular asymmetry may have preceded biochemistry (see Pizzarello et al. 2006 for a recent review).

These data appear very appealing from a broader astrobiological perspective, however, Murchison organic compounds also make up a very complex suite of C, H, N, and O-containing molecular species that are found in a large variety of isomeric forms, up to the limit of their solubility and in declining amounts with lengthening chain lengths. This composition is unequivocally heterogeneous, the likely result of random formative processes and suggestive of a fundamental distinction between a-biotic and biological

chemical processes in their paths to molecular complexity (*e.g.*, Pizzarello 2007). This aspect of chemical evolution lead Luisi (2007) to state that the known abiotic compounds formed riding the free ticket of thermodynamic control and could have carried no opportunities for the inception of biochemistry (*paraphrasis*).

In fact, the analyses of Murchison left many questions unanswered about the potential of its organics for further evolution. Wouldn't the collection of so many compounds, over one thousand, be too much of a good thing? How far would such a "soup" be from reaching the molecular specificity of biochemistry? What could the selective evolutionary factor(s) be? How important was molecular symmetry breaking in chemical evolution? Because CC meteorites collected from falls are rare (just 18 since 1806), Murchison composition is the most studied of all meteorites', and also the most abundant in organic molecular species, the above questions has been readily extended by many to a-biotic chemical capabilities in general.

Indeed, it is as easy to imagine the existence of a large variety of cosmochemical regimes and the likelihood of largely different synthetic outcomes. This was proven recently by the analyses of a different family of CC, the Renazzo type (CR), collected in the Antarctica fields. These meteorites have shown an organic composition unlike any other studied so far, where water soluble compounds are preponderant and, between them, N-containing ammonia, amino acids and amines dominate in abundance the O-containing species such as carboxylic and dicarboxylic acids (Pizzarello et al. 2008; Pizzarello and Holmes 2009). In addition, instead of declining abundances with chain length, CR compounds have a large predominance of lower homologs, *e.g.*, glycine, alanine and α-amino isobutyric acid (aiba). As for CMs, the formative origin of the CR organic suite can be traced back to far away cosmic environments, as it was shown by the +7,200 δD value determined for aiba, which is the highest ever measured directly for an extraterretrial molecule and falls within the values determined spectroscopically for interstellar molecules. The two CR meteorites we analyzed were also pristine and did not show any of the terrestrial contamination tell tale signs, such as an L-enantiomer excess for the chiral amino acids also found in terrestrial proteins. This was useful for identifying features that could have been considered tainted in meteorites exposed to more temperate terrestrial conditions and allowed to show that the precursor aldehyde to the diastereomer amino acids isoleucine and *allo*isoleucine carried an original S-asymmetry to the meteorite's parent body (Pizzarello et al. 2008).

Even if yet limited, the analyses of CR2 organic materials already allow important conclusions: 1) that CR2 precursor cosmic environments carried a *de facto* selection of organic compounds that are over abundantly water-soluble, N-containing and of low molecular weight; 2) that this is a composition of high prebiotic appeal. These novel findings allow re-proposing the basic exobiological inquiry of whether extraterrestrial organic compounds contributed to molecular evolution on the early Earth and to ask how, upon delivery of organic materials with CR or mixed CR and CM composition, this evolution might have proceeded.

Between the molecular species found in meteorites, it is easy to single out amino acids for possible emergent properties because some of these compounds may have reached the Earth with a selected abundant distribution of small molecular while others could have been non racemic (Cronin and Pizzarello 1997) as well as asymmetric catalysts (Pizzarello and Weber 2004). All, uniquely, might have found their path to peptide formation and made an evolutionary story an emergent story.

There are several relevant facts that we do know about amino acids. These compounds are the components of biopolymers indispensable to extant life and are particularly suited to polymerize, even a-biotically and under early Earth conditions. As early as 1961, Oro' and Gudri showed that glycine, when put in the presence of

ammonia and little water like a CR suite could, readily polymerizes with temperatures around 140°C; Leman, Orgel and Ghadiri (2004) showed that the simple presence of carbonyl sulfide, as around volcanoes, could lead to easy formation of peptides; Toniolo et al. (2006) demonstrated that, upon formation of even short peptides, the non-racemic amino acids of meteorites could drive the formation of homochiral helixes. Furthermore there are innumerous possibilities of interactions between amino acids and early Earth environments, as when Orgel (1998) spoke of and experimented with polymerization on the rocks (*e.g.*, Hill et al. 1996).

How far these incremental steps could have reached on an evolutionary scale is hard to say. If life is difficult to define, so is its origin. However, it is conceivable that a determined pursuit could get to demonstrate the proper energetic and catalytic circumstances by which these and other venues of a-biotic chemistry would mimic some of life's complex functions. Since the definition of these functions is, *per se*, an admission of emergence, we may yet get to imitate life's emergence (or at least elements of it).

References

Cronin JR, Pizzarello S (1997) Enantiomeric excesses in meteoritic amino acids. Science 275:951–955

Eschenmoser A (2008) Fundamental questions: An interview with Albert Eschenmoser. In: The Scripps Research Institute News and Views: 8(13). http://www.scripps.edu/newsandviews/e_20080421/

Hill AR, Bohler C, Orgel LE (1996) Polymerization on the rocks: Negatively-charged alpha-amino acids. Orig Life Evol Biosph 28:235–243

Leman L, Orgel L, Ghadiri MR (2004) Carbonyl sulfide-mediated prebiotic formation of peptides. Science 306:283–286

Luisi PL (2007) Presented at: International School of Complexity-4th Course: Basic questions on the Origin of Life; Ettore Majorana Foundation and Centre for Scientific Culture, Erice, Italy, 1–6 October 2006

Orgel LE (1998) Polymerization on the rocks: Theoretical introduction. Orig Life Evol Biosph 28:227–234

Oró J, Guidry GL (1961) Direct synthesis of polypeptides I. Polycondensation of glycine in aqueous ammonia. Arch Biochem Biophys 93:166–171

Pizzarello S, Weber AL (2004) Meteoritic amino acids as asymmetric catalysts. Science 303:1151

Pizzarello S, Holmes W (2009) Nitrogen-containing compounds in two CR2 meteorites: ^{15}N composition, molecular distribution and precursor molecules. Geochim Cosmochim Acta 73:2150–2162

Pizzarello S, Cooper GW, Flynn GL (2006) The nature and distribution of the organic material in carbonaceous chondrites and interplanetary dust particles. In: Meteorites and the Early Solar System II, Lauretta D, McSween HY (ed) Tucson, AZ: University of Arizona Press, pp. 625–51

Pizzarello S, Huang Y, Alexandre MR (2008) Molecular asymmetry in extraterrestrial chemistry: Insights from a pristine meteorite. PNAS 105:3700–3704

Toniolo C, Crisma M, Formaggio F, Peggion C (2001) Control of peptide conformation by the Thorpe-Ingold effect (C^{α}-tetrasubstitution). Biopolymers 60:396–419

E-mail: pizzar@asu.edu

On Emergence

Stuart Kauffman

Institute for Biocomplexity and Informatics, University of Calgary, Canada

Keywords Emergence and Game of Life • Collectively Autocatalytic Sets • Catalytic Closure • Adaptive Opportunities • Efficient Causes as Deduction

I believe it is useful to distinguish at least three senses of emergence: (i). Are there properties of the "whole" system which are not deducible from their parts plus the local interactions of the parts? There are numerous examples of this type of emergence. One consists in mapping a universal Turing machine onto Conway's Game of Life. The Game of Life is a 2 dimensional cellular automaton with the same "physics" or Boolean rule, realized by each square on the lattice, which may be infinite dimensional. Patterns such as gliders and glider guns, describable at a NEW LEVEL of DESCRIPTION that needs no further reference to the Boolean "physics", are then used to construct the universal Turing machine. On an infinite 2 dimensional lattice, the halting problem arises, so many behaviors of the lattice cannot be deduced from the local Boolean rules. This is a form of emergence. The same property has recently been discovered mapping universal Turing machines onto infinite 2 dimensional Ising models. These proofs however, rely on the infinity of the 2 dimensional lattice, and it is not clear what may be "hiding" in this infinity. (ii) The emergence of collectively autocatalytic sets, as in the theory I have developed and described in the abstract for the question on the origin of catalytic cycles, depends upon primitives called molecules, reactions, and catalysis. At a sufficient diversity the ratio of reactions to molecules is sufficiently high, given a probability that a molecule catalyzes a randomly chosen reaction, collectively autocatalytic sets emerge. This is emergence because the whole collectively autocatalytic set achieves catalytic closure and can reproduce or maintain itself given "food". None of its parts have this property. Catalytic closure is an emergent property. Furthermore, the emergence of the collectively autocatalytic set is an example of Robert Laughlin's "laws of organization", and is notable in that it is not reducible to any specific underlying physics. Indeed, the laws of physics might be "slightly" different and autocatalytic sets might still arise, for example if the constants of nature in General Relativity and the Standard Model were slightly different. Thus, it seems that the theory of autocatalytic sets is not reducible to physics. The theory is a mathematical theory independent of the details of the underlying physics. Of course, entities that instantiate "molecules", "reactions", and "catalysis" are needed. But the theory is mathematical, and the law of organization is mathematical, and can be realized by multiple physical platforms. Thus it is independent of physical laws, and not reducible to them. (iii) The EVOLUTION of autocatalytic sets occurs in a non-ergodic universe which will never make all possible proteins length, e.g. 200. We are on a unique diachronic trajectory. History enters when the space of the possible vastly exceeds what can happen. Autocatalytic sets can evolve and co-evolve by natural selection and make novel niches with one another that depend upon their chemical and physical properties. Thus they are capable of Darwinian pre-adaptations which we cannot pre-state. It follows that we cannot make probability statements about their evolution—indeed the biosphere's evolution of collectively autocatalytic cells—nor are their sufficient natural laws to describe that evolution, as I discuss in detail in "Investigations" and "Reinventing the Sacred". But the lack of sufficient natural laws means that the "becoming" of a biosphere of living systems is both partially non-lawful, hence emergent with respect to physics, and also non-random. What is selected cannot always be pre-stated, but succeeds in the ever changing contexts of opportunities for adaptations which in turn

change the opportunities for adaptations. In this sense, evolutionary theory cannot be epistemologically closed, hence the evolution of life again is ever emergent.

There is a further, and I believe deep issue here. Our overriding scientific world view remains reductionism. This has its roots in Aristotle and Newton. Recall that Aristotle had four "causes", formal, final, material, and efficient. The formal cause of your house is the design for the house. The final, "teleological" cause of the house is your decision to have it built. The material cause(s) are the bricks and other materials used in its construction. The efficient cause (s) are the actual processes of building the house. But Aristotle also suggested a model of scientific explanation: the syllogism and deduction. All men are mortal, Socrates is a man, therefore Socrates is a mortal. Feel the logical force here. With Newton's laws in differential equation form, initial and boundary conditions of billiard balls on a table, we compute the forward trajectory of the balls by integrating the equations. But this integration is precisely Aristotle's "deduction". As Robert Rosen pointed out in "Life Itself", with Newton, we abandoned Aristotle's other causes and kept only efficient cause. And we mathematized efficient cause as deduction.

Once these two steps were taken, scientific explanation was always to be in terms of deductions from efficient causes for every event. So reductionism snaps into place: There MUST be a theory down there which logically entails all that happens in the unfolding of the universe. Stephen Weinberg emailed me that this is the view he holds.

But is this view right? I suspect it is not, and bears on emergence. Philosopher David Depew at a recent conference pointed out that the achievement of an adaptation, say the eye, is "blind teleology". He had in mind just what Dawkins had in mind with "The blind watchmaker", that is, Darwin gave us, via natural selection, a means to achieve the appearance of design without a designer.

Now I ask a new question. Can we talk about "the opportunity for an adaptation", say a red spot that is light sensitive on an offspring of a creature with no red spot. Let me formalize an opportunity for an adaptation as, "The opportunity is possible. It may or may not occur. If it occurs it will TEND to be selected and fixed in the population". (I thank Gordon Kaufman for this formulation.) Now, 'tend' is a dispositional term. The actual events that lead to the fixation of the adaptation in the population are efficient causes. However, can we prestate the necessary and sufficient efficient causes that lead to this fixation of the red spot? In general we have no idea what those events must be, and cannot prestate the necessary and sufficient efficient cause conditions. But this means that we can have NO EFFICIENT CAUSE LAW for the fixation of the adaptation.

Next, what kind of a "cause" is the opportunity for an adaptation. Most importantly, is it an efficient cause, like billiard balls hitting one another on the billiard table? I don't see any sense in which the opportunity for an adaptation is an efficient cause, even if we have no precise definition of an efficient cause. In keeping with Depew's Blind Teleology, I want to say that the opportunity for an adaptation is not an efficient cause, but, thanks to Darwin, a Blind Final Cause. If this is acceptable, then it is no longer true that all that unfolds in the becoming of the universe is due to the unrolling of purely efficient causes! Rather, the blind final cause of an opportunity for an adaptation, if the adaptation is achieved, is not due to an efficient cause, yet changes the future course of evolution of the biosphere. But then, we are free from a 2,500 year old, or 300 year old spell stating that all that occurs in the becoming of the universe is due to efficient cause. And we are free of the view that all that occurs is to be captured by mathematicizing efficient cause as mathematical deduction from "the laws". We are free from the stricture of entailment from some final theory, pace Weinberg. The becoming of the universe, biosphere, human economy, human culture, is broader than this ancient view. It is deeply emergent.

On this view, in contrast to Weinberg's final theory that tautologically entails all that does and can occur, the universe is OPEN. This is a huge shift in our understanding of how the universe becomes.

I end by noting that it may be that our incapacity to prestate Darwinian preadaptations, hence absence of sufficient natu, if thal law, may be merely epistemological. Perhaps there is a law but we cannot know it. By contrast, my critique of efficient cause as sufficient, and the case for blind final cause, is ontological, not epistemological. The universe is ontologically open if I am right.

E-mail: skauff@telus.net

Emergence of Self-Movement as a Precursor to Darwinian Evolution

Martin Hanczyc[1], Takashi Ikegami[2]

[1]Dept. of Physics and Chemistry, University of Southern Denmark, Denmark; [2]Dept. of General Systems Sciences, University of Tokyo, Japan

Keywords Autopoiesis • Evolution • Homeodynamics • Movement

A deep understanding of the origin of life requires a careful evaluation of terminology and clarification of objectives, notwithstanding the difficulties of defining life itself. A conceptual framework including time, material and process that helps to explain the origin of life will also help us arrive at a definition of life. Our interest lies in how life-like processes can emerge in a collection of molecules. We specifically focus on self-movement as a characteristic of collective matter that is necessary for the emergence of life. Self-movement of collective matter allows for several emergent properties that are necessary for life namely, purposeful behavior, homeodynamics, autopoiesis, primitive cognition, and robustness. We envision that self-movement may be a readily accessible process for diverse collections of matter and that primitive self-moving agents able to sense the environment and move with purpose would constitute the first examples of life on earth. We present a few examples of simple chemical systems that self-organize to produce oil droplets capable of movement, environment remodeling, spontaneous division and chemotaxis. These chemical agents are powered by an internal chemical reaction based on the hydrolysis of an oleic anhydride precursor or on the hydrolysis of HCN polymer, a tentatively plausible prebiotic source of energy and material for the origin of life. Such motile agents would be capable of competing for resources, escaping dangers, sustaining themselves while at the same time retaining a chemical memory of their past actions. Such a system would foster the development of more complex internal chemical networks which would be responsible for self-maintenance, self-movement and resource allocation/exploitation. In this sense these primitive agents would be capable of temporal evolution.

Many physicists and chemists believe that living states are realization of dissipative structures; nonlinear systems in general generate organized structures under the non-equilibrium open flow conditions. Examples exist from Benard cells to economic market organizations. A problem is how the non-equilibrium open flow state can be self-sustained by the system. Here our oil droplet sustains the condition by self-movement through the coupling between chemical reaction and the internal convection flow.

True Darwinian evolution as an emergent property of the system would come later in the timeline of early Earth when the best self-moving agents containing chemical networks acquire the ability to reproduce. We should discuss not only our true ancestor but also other

candidates that did not directly contribute to our living systems in order to understand why some material systems could not become life. Therefore we believe that it is essential to study the role of self-movement in the origin of life as an early state of matter that precedes replication in order to understand the transition from non-living but self sustaining chemical networks to evolving cellular life.

References

Hanczyc M, et al (2007) J Am Chem Soc 129:9386
Ikegami T, Suzuki K (2008) BioSys 91:388

E-mail: martin@ifk.sdu.dk; ikeg@sacral.c.u-tokyo.ac.jp

Emerging Properties in the Origins of Life and Darwinian Evolution

Andrew Pohorille

Exobiology Branch, NASA Ames Research Center, USA

Keywords Emergent Properties • Darwinian Evolution • Vesicles • Protocells

It would be surprising to find respectable scientists who still hold a "vitalistic" view of living systems as being governed by different principles than non-living ones. It is firmly believed that principles of physics and chemistry are sufficient to describe how life emerged from non-living parts. This does not, however, mean that the point is moot. Instead it can be rephrased—what are the essential emergent properties of life, what processes bring them to existence and how do they differ from analogous processes acting on non-living systems?

In both living and non-living systems emergent properties arise when a number of simple agents (parts of the system) generate complex behavior that cannot be easily predicted from properties of these agents. There are at least four types of emergent traits that are particularly relevant to the origins of cellular life:

1. organic/synthetic chemistry,
2. macromolecular and sub-cellular structures and functions,
3. metabolic and regulatory networks,
4. cellular behavior

In many instances, these emergent traits refer to different spatial and temporal scales. This is in accord with our understanding of emergent behavior as often arising from coupling between interactions at different levels.

As an example, macromolecular and sub-cellular structures and functions exhibit typical characteristics of emergent phenomena. It would be, for example, difficult to anticipate the existence of vesicles from observing single amphiphilic, membrane-forming molecules. Similarly, it would be difficult to predict the structure and function of ribosomes or energy transduction systems only from the knowledge of each individual component. Nevertheless, considerable progress has been made in understanding how these and many other macromolecular systems came to existence. Tracing how this progress has been achieved sheds light on the key processes that lead to emergent behavior in biological systems and

how they differ from analogous processes in non-biological systems, especially those that are products of human activity.

In everyday life we constantly encounter emergent systems. A watch, a car or a computer could serve as examples. All of them are products of designs aimed at achieving certain goals that are built on purpose according to specific blueprints. In contrast, in biological systems there are no goals, designs, purposes and blueprints. Any explanation of emergent properties in living systems that directly or indirectly refers to these concepts should be rejected as invalid. Instead, emergent properties are products of interplay between principles of physics and chemistry and Darwinian evolution. This means that emergent behavior is just another evolutionary trait, which can be and should be considered in the framework of the theory of evolution. It also implies that our primary focus should be on emerging functions rather than structures because functions and not structures are the basis for Darwinian selection.

The relation between emergent properties and evolution in the origins of life can be well illustrated in the example of vesicles, which form membranous compartments composed of amphiphilic molecules (Pohorille and Deamer 2009). Vesicles can be considered typical emergent structures held together by forces of self-assembly. It is frequently thought that their main role in the origin of cellular life was to provide containment so that nascent cells would not loose essential internal components to the surrounding environment. This view is greatly simplified. Membranes mediated a number of essential functions, such as delivery of nutrients, catalysis of chemical reactions, capture of energy and its transduction into the form usable in chemical reactions, transmission of environmental signals, cellular growth and cell volume regulation. This means that their evolution was closely coupled with evolution of other emergent properties of protocells and understanding this coupling can provide important clues to the origins of life. Specifically, the earliest protocellular membranes consisted, most likely, of simple amphiphiles, such as fatty acids and their glycerol esters (Hargreaves and Deamer 1978; Walde et al. 1994; Hanczyc et al. 2003). Such membranes are relatively permeable to ionic species, including protons. This implies that maintaining proton gradients across the earliest cell walls would have been quite difficult, from which it follows that such gradients were not initially used for energy storage and utilization. Only as evolution progressed to yield less permeable, lipid membranes, transport of ions became more difficult and required assistance (Nozaki and Tanford 1981; Paula et al. 1996; Pohorille and Deamer 2009). This gave raise to another emergent structures—ion channels, which are self-assembling associations of small proteins lodged in membranes that mediate ion transport. Some very simple proteins are capable of forming highly efficient channels, which is a property difficult to predict from their sequence or structure (Pohorille et al. 2005). An important consequence of forming less permeable membranes was the ability to maintain proton gradients. This, in turn, made possible the emergence of primitive bioenergetic functions, in which energy captured from the environment was stored as proton gradients and subsequently utilized to synthesize energy-carrying molecules (Deamer 1997). This evolutionary advance facilitated the emergence of molecular devices capable of active transport whereby energy was used to move ions and other molecules against their concentration gradients. These two landmark innovations turned protocells into true molecular machines that used external resources to do work in ways that were previously impossible. Even today understanding how all components of these machines work in concert and designing even the simplest models of protocells in the laboratory remains a challenge.

Darwinian evolution forms a very powerful paradigm for studying emergent behavior in biological systems. It also means that a number of difficulties encountered in theory of

evolution inevitably propagate to investigations of emergent properties. One of them is to provide general explanation how small mutations accumulate to yield qualitative changes often associated with emergent behavior. It also has methodological implications for studies of the origins of life. Specifically, we have considerable successes in explaining evolution. In contrast, we have been remarkably unsuccessful in predicting evolutionary changes. From this observation it follows that it might be advantageous to reason about emerging properties starting from modern systems and extrapolating backwards in time to protocellular systems. However, for some emergent properties this is not possible because the evolutionary gap is too large to make such extrapolation meaningful. Then we are forced to work in a less certain, inductive framework. Another methodological problem is that emergent, supposedly protobiological systems are often specifically designed in a laboratory for a given purpose and might not lie on or illustrate any realistic evolutionary pathway. Then, their relevance to the origins of life is questionable.

References

Deamer DW (1997) The first living systems: a bioenergetic perspective. Microbiol Mol Biol Rev 61:239–261

Hanczyc MM, Fujikawa SM, Szostak JW (2003) Experimental models of primitative cellular compartments: encapsulation, growth, and division. Science 302:618–622

Hargreaves WR, Tanford C (1981) Liposomes from ionic, single-chain amphiphiles. Biochemistry 17:3759–3768

Nozaki Y, Tanford C (1981) Proton and hydroxide ion permeability of phospholipid vesicles. Proc Natl Acad Sci USA 78:4324–4328

Paula S, Volkov AG, Van Hoek AN, Haines TH, Deamer DW (1996) Permeation of protons, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness. Biophys J 70:339–348

Pohorille A, Deamer DW (2009) Self-assembly and function of primitive cell membranes. Res Microbiol doi:10.1016/j.resmic.2009.06.004

Pohorille A, Schweighofer K, Wilson MA (2005) The origin and early evolution of membrane channels. Astrobiology 5:1–17

Walde P, Wick R, Fresta M, Mangone A, Luisi PL (1994) Autopoietic self-reproduction of fatty acid vesicles. J Am Chem Soc 116:11649–11654

E-mail: Andrew.Pohorille@nasa.gov

On Question 2: Emergence

Víctor Norris, Yohann Grondin

AMMIS Laboratory EA 3829, University of Rouen, France

Keywords Molecular Complementarity • Evolution • Emergence • Hyperstructure • Coherence • Competition

Introduction

Emergent properties, which are often considered as central to the nature of Life, resist attempts to predict or deduce them (Van Regenmortel 2004). In pursuit of the bases of a

unifying theory of biology, we have developed the concept of *competitive coherence* in which emergent properties arise. This concept, which is essentially biological, describes the operation of the many organisations that are constrained by the need to reconcile coherence with their present environment (both internal and external) and coherence with their past environments. Competitive coherence played a major role in the 'ecosystems first' approach to the origins of life in which an ecosystem of non-covalent assemblies of molecules or *composomes* (Segre, Ben-Eli et al. 2000) swapped constituents in an abiotic flow of creation and degradation (Hunding, Kepes et al. 2006; Norris, Hunding et al. 2007; Root-Bernstein 2009). In our hypothesis, the composition of these composomes progressively changed as a result of molecular complementarity which protected abiotically created molecules from degradation (Root-Bernstein and Dillon 1997). Although it is difficult to predict exactly which molecules and which complementary structures will be important in a system, the concept of competitive coherence helps explain how and why biological systems, molecular complementarity and emergence are related.

Competitive coherence

Biological systems on all scales are often compelled to obtain a future state that is coherent with environmental conditions and with previous states. These states are created by the combined functioning or activity of a set of constituents of the system. This active set is selected from the larger set available to the system. To grow and survive, bacteria for example must select an active set of molecules and macromolecules in response to external and internal conditions and to their history. Such responses entail (1) the generation of a coherent cell state, in which the cell's contents work together efficiently and harmoniously with one another and with the environment, and (2) the generation of a coherent sequence of cell states. Contradiction and incoherence are punished since, for example, a cell that simultaneously induces the expression of genes for growth at high temperature and at low temperature is likely to be out-competed by rival cells that induce each set of genes only when needed. A cell that goes from one cell state to another very different state (without good environmental reason) risks wasting precious resources. A strong selective pressure therefore exists to generate active sets of constituents to provide both coherent cell states and a coherent sequence of such states. We have proposed that competitive coherence is responsible for generating these active sets (Norris 1998). This concept is based on the way a system must maintain both the continuity of the composition of its active set via a Next process and the coherence of this active set (with respect to the inside and outside world) via a Now process. In one in silico implementation of this concept, the state of a system at time n+1 is determined by a competition between the Next process, which is based on its state at time n, and the Now process, which is based on the developing n+1 state itself (Norris, unpublished). In the case of amateur football, consider the problem of selecting the team (the active set) each week from a larger group of potential players. A Next process might be the tendency for someone who plays this week to be someone who plays next week (it is, for example, easier to arrange shared travel with those already present). A Now process might be the coherence of the team with respect to itself (the team must have a goalkeeper, defenders and attackers) and with respect to its opponents (who might be particularly brutal). As the new team is being chosen, the Next process gives way in importance to the Now process. Competitive coherence also operates at the higher level of the football league itself: a Next process results in teams that are in the league in one season being likely to be in it the next season whilst a Now process results in a coherent league with teams of the same level dispersed over a certain geographical region.

Competitive coherence in bacteria

RNA polymerase is limiting for the transcription of genes in the bacterium *Escherichia coli* so many genes cannot participate in the active set (Stickle, Vossen et al. 1994) and ribosomes are limiting for the translation of mRNA (Vind, Sorenson et al. 1993). More recently, it has been proposed that the use of reaction networks in bacteria is also limited to a small active set (Nishikawa, Gulbahce et al. 2008). In the case of bacterial organisation, competitive coherence selects a particular hyperstructure within a bacterium (such hyperstructures include multi-macromolecule assemblies responsible for ribosome synthesis, chemotaxis, sugar metabolism, DNA replication, RNA turnover etc. (Norris, den Blaauwen et al. 2007)); a Next process allows those genes that are already expressed as part of a hyperstructure to help determine which genes are expressed in the next time step in a hyperstructure; a Now process then allows those genes that are starting to be expressed together in a hyperstructure to recruit related genes to the hyperstructure. Many of these hyperstructures fall into the *functioning-dependent* class of structures that form when their constituents are performing a task and that dissociate when this task is over (for references see (Thellier, Legent et al. 2006)).

Competitive coherence operates at the higher level of the bacterial cell itself such that the state of a cell at any one time corresponds to the set of hyperstructures present within it (Norris, Blaauwen et al. 2007). A new cell state is the result of (1) a Next process whereby the current active set of hyperstructures in the cell determines the next active set and (2) a Now process whereby the developing set of hyperstructures progressively recruits, maintains or dismisses hyperstructures. Competition between these two processes ensures a sequence of sets of hyperstructures (cell states) that optimises growth and/or survival.

Finally, competitive coherence operates at the level of colonies and biofilms in which different species of bacteria compete and collaborate with one another. For example, there are over 500 taxa of oral bacteria and, in dental plaque, the coaggregating partnership of *Streptococcus oralis* and *Actinomyces naeslundii* may allow each to grow where neither can grow alone (Rickard, Gilbert et al. 2003).

Emergence and competitive coherence

In the framework of competitive coherence, emergence is related to the selection of the subset of constituents that are active together (Norris, Cabin et al. 2005). Suppose each constituent has a large number of characteristics (as in the case of macromolecules such as mRNA and proteins which contain a large number of sites that can bind water, ions, molecules and other macromolecules). As proteins are being chosen via competitive coherence to work together, suppose that the first ones to be chosen just happen to contain a binding site to the same molecule. Suppose that, in some environments, this combination of proteins proves useful. Suppose too that this molecule becomes available, perhaps for the first time. The presence of this binding site could then become an important factor in the coherence process which dominates the choice of the rest of the proteins to work together in the active set. In other words, the environment acts via the coherence process to lend importance to one out of many sites. The result is the selection of this site (plus the molecule that binds to it) as a determinant of the cell's response to a particular environment. More specifically, consider, for example, that (1) this binding site is for a particular phospholipid with long, saturated acyl chains and (2) the proteins with this site bind to the phospholipid to form a domain in which they are juxtaposed and in which their activities complement one another. There might then be a selection for this binding site in other complementary proteins. In the language of competitive coherence, binding to this phospholipid would become a type of connectivity to determine membership of an active

set and this active set would take on the physical form of a proteolipid domain responsible for a particular function. Hence emergence in the context of competitive coherence can be understood in terms of a new criterion for membership of the active set. Possible examples of such criteria include DNA supercoiling, ion condensation, affinities for molecules and structures, and convergence on a common vibrational mode (Norris, Blaauwen et al. 2007). In the context of competitive coherence and the origins of life, the most important of the determinants of membership of the active set was molecular complementarity.

Molecular complementarity

The basis of receptor evolution has been attributed to small molecule homo- or self-complementarity in that receptors often contain in their binding regions copies of their own ligands (Dwyer 1989; Dwyer 1998). Such receptors would be complementary not only to their ligands, but potentially to each other. This theory has been developed to incorporate heterocomplementarity so that small molecules that are complementary to each other help each other's receptors evolve (Root-Bernstein 2005). For example, insulin is homocomplementary as well as heterocomplementary to glucagon, and copies of insulin are found in the binding regions of both the insulin and glucagon receptors. Similarly, insulin binds glucose, and multiple copies of insulin make up the core of the glucose transporter.

In the flux of creation and degradation that characterised the prebiotic world, species of molecules that interacted were preserved as composomes (Segre, Ben-Eli et al. 2000; Hunding, Kepes et al. 2006; Norris, Hunding et al. 2007; Root-Bernstein 2009). These interactions were based on molecular complementarity. Hence, from the competitive coherence standpoint, the composomes were the active sets and molecular complementarity was the criterion which led to the emergence of these active sets.

Discussion

Life is about behaviour and the fundamental basis of behaviour, we argue, is the process of competitive coherence which describes the way biological systems achieve coherence with both their history and their environment. Not surprisingly therefore, we have proposed that the concept may help in the search for quantitative basis for biological complexity (Norris, Cabin et al. 2005). Competitive coherence is related to concepts such as synergies (Haken 1983; Kelso 2008), SOWAWN machines (Ji 2009) and neural Darwinism (Edelman 1987). In artificial learning experiments to test competitive coherence, active sets can become selected due to common features in patterns of connectivity (Norris, unpublished). The possession of these common features constitutes emergence but there is nothing magical or vitalistic about this emergence or the competitive coherence which produces it.

Competitive coherence is a scale-free concept that operates at levels ranging from macromolecular assemblies to social groups. The nature of the Next and Now processes varies with the level. At the level of bacteria, these processes take the form of site-binding, DNA supercoiling, transcription factors, ion condensation etc. At the level of human groups, these processes include the mafia and *status quo* pressures that are familiar to us all. Many social organisations are constrained by the need to reconcile coherence with their present environment (both internal and external) and coherence with their past environments. To grow and survive, research laboratories, for example, have to select an active set of workers in response to new discoveries and to new funding initiatives but must reconcile this selection with the research history of the laboratory and, in particular, with its skills, experience and interests.

Could competitive coherence operate at higher levels still? And if it were to operate, what form would the Now and Next processes take and what would emerge? One candidate is subjective experience, the role of which remains controversial. A role for subjective

experience in Now and Next processes at the level of ecosystems would support speculations that our world itself is 'conscious' (Norris 1998).

References

Dwyer DS (1989) Amino acid sequence homology between ligands and their receptors: potential identification of binding sites. Life Sci 45(5):421–9

Dwyer DS (1998) Assembly of exons from unitary transposable genetic elements: implications for the evolution of protein-protein interactions. J Theor Biol 194(1):11–27

Edelman G (1987) Neural Darwinism. The Theory of Neuronal Group Selection. New York, Basic Books

Haken H (1983) Synergetics, an introduction: Nonequilibrium phase transitions and self-organization in physics, chemistry and biology. Berlin, Springer

Hunding A, Kepes A et al (2006) Compositional complementarity and prebiotic ecology in the origin of life. Bioessays 28(4):399–412

Ji S (2009) Molecular theories of the living cell: conceptual foundations, molecular mechanisms and applications. New York, Springer

Kelso JAS (2008) Synergies: Atoms of brain and behavior. A multidisciplinary approach to motor control. D. Sternad. Heidelberg, Springer

Nishikawa T, Gulbahce N et al (2008) Spontaneous reaction silencing in metabolic optimization. PLoS Comput Biol 4(12):e1000236

Norris V (1998) Bacteria as tools for studies of consciousness. Toward a Science of Consciousness II: The Second Tucson Discussions and Debates. S. Hameroff, A. Kaszniak and A. Scott. Cambridge, USA, MIT Press: 397–405

Norris V (1998) Modelling *E. coli*: the concept of competitive coherence. Comptes Rendus de l'Academie des Sciences 321:777–787

Norris V Blaauwen TD et al (2007) Toward a Hyperstructure Taxonomy. Annu Rev Microbiol 61:309–329

Norris V, Cabin A et al (2005) Hypercomplexity. Acta Biotheor 53(4):313–30

Norris V, den Blaauwen T et al (2007) Functional taxonomy of bacterial hyperstructures. Microbiology and Molecular Biology Reviews 71(1):230–53

Norris V, Hunding A et al (2007) Question 7: the first units of life were not simple cells. Orig Life Evol Biosph 37(4–5):429–32

Rickard AH, Gilbert P et al (2003) Bacterial coaggregation: an integral process in the development of multi-species biofilms. Trends Microbiol 11(2):94–100

Root-Bernstein RS (2005) Peptide self-aggregation and peptide complementarity as bases for the evolution of peptide receptors: a review. J Mol Recognit 18(1):40–9

Root-Bernstein RS (2009) An ecosystems-first theory of the origin of life based on molecular complementarity. Astrobiology: Emergence, Search, and Detection of Life. V. A. Basiuk, American Scientific Publishers: 1–30

Root-Bernstein RS Dillon PF (1997) Molecular complementarity I: the complementarity theory of the origin and evolution of life. J Theor Biol 188(4):447–79

Segre D, Ben-Eli D et al (2000) Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. Proceedings of the National Academy of Science USA 97:4112–4117

Stickle DF, Vossen KM et al (1994) Free DNA concentration in *E. coli* estimated by an analysis of competition for DNA binding proteins. Journal of Theoretical Biology 168:1–12

Thellier M, Legent G et al (2006) Steady-state kinetic behaviour of functioning-dependent structures. The FEBS journal 273(18):4287–99

Van Regenmortel MHV (2004) Emergence in Biology. Modelling and simulation of biological processes in the context of genomics. P. Amar, J.-P. Comet, F. Kepes and V. Norris. Evry, France, Genopole: 123–132

Vind J, Sorenson MA et al (1993) Synthesis of proteins in *Escherichia coli* is limited by the concentration of free ribosomes. Expression from reporter genes does not always reflect functional mRNA levels. Journal of Molecular Biology 231:678–688

E-mail: victor.norris@wanadoo.fr

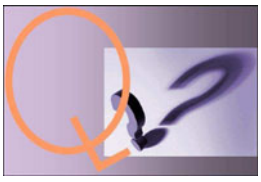
Workshop OQOL'09

Extended Abstracts for the Following Selected Question

- **Heterotrophic Versus Autotrophic Scenarios**

Premise. One of the important questions relating to the origin of life problem today is the *heterotrophy/autotrophy* dichotomy. In an (extreme) heterotrophic scenario, the organic material supposed to have accumulated in a prebiotic world by high-energy processes (such as those of the Miller type in a primordial atmosphere, or by impact delivery to the Earth from extraterrestrial sources) is assumed to generate the critical self-organization processes culminating in life's origin. In sharp contrast, in an (extreme) autotrophic scenario, this kind of organic material is considered irrelevant and it is, instead, postulated that the substrates and intermediates of the chemical processes that organized themselves toward life were generated through synthetic processes within self-organized structures (e.g. from free-energy reach C-1- or C-2-organics, combined with strong inorganic reductants).

The question. Do you see strong chemical arguments in favor of the one or the other scenario? And which experiments would you do/suggest, in order to possibly clarify this dichotomy?



The Combination of Chemical, Biochemical, and Geological Arguments Favours Heterotrophic Scenarios

Robert Pascal

Institut des Biomolécules Max Mousseron, University of Montpellier 2, France

Keywords Coupled Reactions • Early Earth Atmosphere • Energy Carriers • Heterotrophy • Hydrogen Escape

The independent proposal by Oparin (1924) and Haldane (1928) that early living organisms lived at the expense of abiotically synthesized organics (as sources of both organic matter and energy) constituted a major conceptual breakthrough in the study of the origin of life. It founded the basis of an experimental approach. But the possibility that the primitive Earth was not favourable to synthesis led to the alternative hypothesis that inorganic sources of energy induced the self-organization of metabolic pathways (Wächtershäser, 1988). What actually occurred depends on either the environment of the primitive Earth was capable of providing high-energy organics in quantity sufficient to initiate self-organization, or geological processes allowed inorganic redox components to be brought into contact and to deliver the free energy needed to drive a metabolism. Analyzing the efficiency of the processes involved in both scenarios in relation with the requirements for the first living systems may be helpful in solving this alternative.

The Early Earth Atmosphere, Redox Gradient, and Abiotic Synthesis

The idea that a redox gradient between the mantle and the surface constitute a source of free energy for early living organisms in specific locations where strong inorganic reducing agents come into contact with the surface as a consequence of geodynamics is attractive. This issue is connected to that of the composition and evolution of the early atmosphere since processes that oxidize the atmosphere tend to increase the redox gradient with the mantle. The widespread belief was that a fast escape of hydrogen atoms to the outer space (Kasting, 1993) drove the atmosphere of the Earth towards a mixture of CO_2 and N_2 as major constituents. However, in a CO_2-rich and O_2-free upper atmosphere, the temperature was much lower and the photolysis of hydrogen-containing molecules (H_2O, CH_4…) less efficient to promote the escape of hydrogen atoms (Tian et al. 2005). Unless another process could substitute for hydrogen escape, the volcanic emission of reduced species was not compensated to maintain a neutral composition. Redox gradients with highly reducing minerals would have been less potent as energy sources to feed the metabolism of early living organisms. But, this is not incompatible with the formation of substantial amounts of inactivated organic derivatives by heating dissolved mixture of inorganic precursors in hydrothermal systems. Moreover, UV-irradiation, lightning, or impacts promoted the formation of activated organic compounds in reducing, or even neutral (Cleaves et al. 2008), atmospheres and delivered them to the surface, supporting a heterotrophic scenario. Among these compounds, energy-rich low-molecular weight organics are prone to undergo further synthetic processes provided that they are concentrated in specific locations on the surface rather than diluted in the ocean. In this view, the question of autotrophy or heterotrophy can only be handled by considering both the chemistry derived from potential energy sources and the geological processes capable of segregating inorganic or organic chemicals in favourable locations.

Self-organization Requires Activated Molecules

The availability of high-energy organic carriers was essential to the development of the metabolism of early organisms (Lazcano and Miller 1999). Actually, all forms of energy are not equivalent in inducing self-organization. The Second Law of thermodynamics, implying that energy flows from high-energy intermediates to less concentrated forms, requires these carriers, as ATP in biology, to deliver energy through coupled reactions in amounts per chemical event sufficient to convert endergonic processes into thermodynamically favourable ones. These processes may have triggered the development of metabolic cycles (Shapiro 2006) beginning a self-organization process. However, supporters of a

genetic-first process (Pross 2004; Orgel 2008) have challenged the view that metabolic networks could initiate biological self-organization. Anyway, even the replication of genetic polymers requires activated monomers and then the supply of energy sufficient in quantity and quality to drive the process. Moreover, early living organisms were capable of transforming a part only of the available sources of energy and had to rely on the strongest ones. The fact that translation is one of the most conserved processes supposes that any scenario of early evolution must account for the formation of the amino acid activation intermediates (Pascal et al. 2005) and especially aminoacyl adenylates ($\Delta G^{\circ\prime}=-70$ kJ mol^{-1}) that are highly unstable (Wells and Fersht 1986). Inorganic redox potential proposed to have played a role in an autotrophic scenario are in principle unable to drive the formation of those essential biochemical intermediates. For instance, a value of −31.5 kJ mol^{-1} has been reported (Schoonen et al. 1999) for the standard free energy for the formation of pyrite (Eq. 1),

$$FeS + H_2S \rightleftarrows FeS_2 + H_2 \quad (1)$$

Moreover, converting redox potentials into free energy available for a metabolism is a difficult task. It requires connecting redox reactions (already resulting from the association of half-reactions) and energy transfers through afterwards coupled reactions, whereas the corresponding processes of electron and chemical group transfers take place on different time scales. The efficient coupling of these two kinds of processes has been achieved by living organisms after chemiosmosis (Mitchell 1961) emerged, but through a highly complex mechanism involving the transient formation of a proton gradient across the membrane.

Conclusion

Requirements on the availability of free energy sources/carriers have been identified, which are dependent on the environment of the primitive Earth and on geodynamic/atmospheric processes. Although these views are fully consistent with the formation of organic compounds in hydrothermal systems (another heterotrophic scenario), the inorganic sources of energy (pyrite formation) are not considered as able to deliver free energy in amounts (per chemical events) sufficient to induce biological self-organization. On the contrary, activated low molecular weight organic compounds with energy contents sufficient to induce the self-organization of metabolic and/or genetic systems are likely to have been formed by photolysis or lightning.

References

Cleaves HJ, Chalmers JH, Lazcano A, Miller SL, Bada JL (2008) A Reassessment of Prebiotic Organic Synthesis in Neutral Planetary Atmospheres. Orig Life Evol Biosph 38:105–115

Haldane JBS (1929) The Origin of Life. The Rationalist Annual. Reprinted in: Bernal JD (1967) The Origin of Life The Weidenfeld and Nicolson Natural History Carrington R (ed) Readers Union, London, p. 242–249

Kasting JF (1993) Earth's Early Atmosphere. Science 259:920–926

Lazcano A, Miller SL (1999) On the Origin of Metabolic Pathways. J Mol Evol 49:424–431.

Mitchell P (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature 191:144–8

Oparin AI (1924) The Origin of Life. Proiskhodenie Zhini. English translation in: Bernal, JD (1967) The Origin of Life The Weidenfeld and Nicolson Natural History Carrington R (ed) Readers Union, London, p. 199–234

Orgel LE (2008) The Implausibility of Metabolic Cycles on the Prebiotic Earth. PLoS Biol 6:e18
Pascal R, Boiteau L, Commeyras A (2005) From the Prebiotic Synthesis of α-Amino Acids Towards a Primitive Translation Apparatus for the Synthesis of Peptides. Top Curr Chem 259:69–122
Pross A (2004) Causation and the origin of life. Metabolism or replication first? Orig Life Evol Biosph 34:307–321
Schoonen MAA, Xu Y, Bebie J (1999) Energetics and kinetics of the prebiotic synthesis of simple organic acids and amino acids with the FeS-H_2S/FeS_2 redox couple as reductant. Orig Life Evol Biosph 29:5–32
Shapiro R (2006) Small molecule interactions were central to the origin of life. Q Rev Biol 81:105–125
Tian F, Toon OB, Pavlov AA, De Sterck H (2005) A Hydrogen-Rich Early Earth Atmosphere. Science 308:1014–1017
Wells TNC, Ho CK, Fersht, AR 1986. Free Energy of Hydrolysis of Tyrosyl Adenylate and Its Binding to Wild-Type and Engineered Mutant Tyrosyl-tRNA Synthetases. Biochemistry 25:6603–6608

E-mail: rpascal@univ-montp2.fr

Experimental Model Protocells Support a Heterotrophic Origin of Life

Sheref S. Mansy

Center for Integrative Biology, University of Trento, Italy

Keywords Heterotrophy • Protocell • Origin of Life • Prebiotic

Life can be divided into two categories, those that acquire their nutrients and those that make their own, i.e. organisms are either heterotrophic or autotrophic. The question is then, which of these two forms of life are simpler and thus more likely to have arisen first? Historically, heterotrophy has been viewed as the more likely starting point (Oparin 2003), due in part to its perceived simplicity. However, simplicity in cellular structure requires greater complexity in the environment, in the form of nutrients, to compensate for the cell's inability to produce its own nutrients. The availability of abiotically synthesized nutrients in the environment is supported by simulated prebiotic syntheses of amino acids (Miller 1953), nucleobases (Oro 1961), nucleotides (Powner et al. 2009), sugars (Ricardo et al. 2004), and lipids (Hargreaves et al. 1977). However, many are unconvinced by such arguments and instead believe that early life could not have survived off of "free lunch" (Morowitz 1992). In other words, cells that are incapable of providing for themselves would have quickly exhausted the nutrients available in their surroundings and died. Therefore, only autotrophic organisms could have survived the environments of early Earth. The two opposing views, not surprisingly, have led to two different modes of research. The heterotroph supporters tend to focus on the creation of self-replicating systems dependent upon provided energy sources, and the autotrophic supporters often focus on geochemical cycles that mimic contemporary biochemical paths. In short, heterotrophy versus autotrophy emerges as replication-first versus metabolism-first arguments.

One strategy to resolve this debate is to try to build model protocells in order to evaluate which paths are compatible with the measured behavior of model protocells of specific

compositions. The approach is attractive in that it attempts to piece together life rather than to characterize cellular components in isolation. However, those that attempt to build protocells often have different views as to what threshold needs to be crossed for life-like properties to emerge. Much of the protocell experiments are aimed at creating self replicating systems (Szostak et al. 2001), which is in contrast with those that contend that life is a self sustained entity that does not necessarily require replication (Luisi 2003). The latter has been coined autopoiesis. Of these two protocell perspectives, the replication model fits more easily within heterotrophy, whereas the autopoiesis model seems better accommodated by autotrophy, although interesting examples of heterotrophic-like autopoiesis exist (Zepik et al. 2001). The allure of heterotrophic, replicating protocells is its apparent simplicity and compatibility with Darwinian evolution. However, the autopoietic perspective may more closely describe what we recognize as life (Luisi 2006).

A simple parameter to evaluate in order to gain insight into the likelihood of heterotrophic and autotrophic origins is membrane permeability. Highly impermeable membranes give compartments sealed off from their surroundings and thus are incapable of acquiring or releasing material. Such a protocell needs to rely on internally synthesized nutrients and thus would be an autotroph. Permeable membranes, conversely, allow for the uptake and release of nutrients and waste. Therefore, such a heterotrophic protocell could survive by acquiring externally supplied nutrients. A series of permeability studies of membranes composed of prebiotically plausible, monoacyl lipids, such as fatty acids, show a high degree of permeability and selectivity (Hargreaves and Deamer 1978; Chen and Szostak 2004b; Sacerdote and Szostak 2005). Fatty acid vesicles even allow for the passage of nucleotides in the absence of specific transport machinery (Walde et al. 1994a; Chen et al. 2005; Mansy et al. 2008). In summary, the permeability properties of model protocell membranes composed of fatty acids are amenable to heterotrophic, but not autotrophic, processes.

Since model protocell membranes are permeable to nucleotides, they can be used to create heterotrophic cell-like structures capable of genetic replication. Remarkably, only four components (fatty acids, primer, template, and activated nucleotides), excluding salts and buffers, are required to generate a system that acquires nutrients, i.e. activated nucleotides, to fuel compartmentalized copying of a nucleic acid template (Mansy et al. 2008). The permeability properties of the system are dependent upon the lipid composition of the membrane. The permeability advantages of fatty acid membranes complement well other properties, such as their ability to easily form boundary structures (Gebicki and Hicks 1973; Hargreaves and Deamer 1978), grow (Berclaz et al. 2001; Hanczyc et al. 2003; Chen and Szostak 2004a), replicate (Walde et al. 1994b; Hanczyc et al. 2003; Zhu and Szostak 2009), and compete for resources (Chen et al. 2004). Fatty acid vesicles are also stable enough to survive temperature fluctuations that melt and anneal duplex nucleic acids (Mansy and Szostak 2008). Prebiotically plausible fatty acid vesicles allow for the emergence of a variety of life-like properties that also commit the resulting cell-like structure to heterotrophic means of survival.

What has been briefly described is a one sided story. Unfortunately, few attempts to generate autotrophic protocells or protocellular systems composed of less permeable diacyl phospholipids have been reported. This is in spite of interesting data that show that light driven reactions in the presence of weak acids can generate pH gradients across diacyl phospholipid membranes (Deamer and Harang 1990; Deamer 1992; Sun and Mauzerall 1996). Diacyl phospholipid systems are also attractive due to their increased stability to a wider variety of conditions. It seems likely that our understanding of protocellular processes would be enriched if more effort were expended in developing such autotrophic systems.

If one were to base their opinions on experimental evidence, then the conclusion drawn likely would be that life had a heterotrophic beginning. This, of course, does not mean that it is correct. Instead, the paucity of data indicates the need for more protocell experiments so that each data set can be better evaluated against competing theories. Nevertheless, the protocell perspective may, to a certain extent, avoid the posed question of the nature of Earth's first cells. One could envisage the appearance of heterotrophic protocellular structures that then evolve into autotrophic cells (rather than protocells). In this case, the evolution of metabolism would likely have been coupled with the evolution of increasingly impermeable membranes (Szathmary 2007). Similarly, focus on the emergence and evolution of a single protocell type may be misguided. Perhaps mixtures of cell-like structures, some of which are better described as autotrophic while others as heterotrophic, is a more realistic view of evolution. One hopes that through experimental research, we will be better able to answer such questions.

Acknowledgments
SSM is supported by the Armenise-Harvard foundation.

References

Berclaz N, Muller M, Walder P, Luisi PL (2001) Growth and transformation of vesicles studied by ferritin labeling and cryotransmission electron microscopy. J Phys Chem B 105:1056–1064

Chen IA, Szostak JW (2004a) A kinetic study of the growth of fatty acid vesicles. Biophys J 87:988–998

Chen IA, Szostak JW (2004b) Membrane growth can generate a transmembrane pH gradient in fatty acid vesicles. Proc Natl Acad Sci USA 101:7965–7970

Chen IA, Roberts RW, Szostak JW (2004) The emergence of competition between model protocells. Science 305:1474–1476

Chen IA, Salehi-Ashtiani K, Szostak JW (2005) RNA catalysis in model protocell vesicles. J Am Chem Soc 127:13213–13219

Deamer DW (1992) Polycyclic aromatic hydrocarbons: primitive pigment systems in the prebiotic environment. Adv Space Res 12:183–189

Deamer DW, Harang E (1990) Light-dependent pH gradients are generated in liposomes containing ferrocyanide. Biosystems 24:1–4

Gebicki JM, Hicks M (1973) Ufasomes are stable particles surrounded by unsaturated fatty acid membranes. Nature 243:232–234

Hanczyc MM, Fujikawa SM, Szostak JW (2003) Experimental models of primitative cellular compartments: encapsulation, growth, and division. Science 302: 618–622

Hargreaves WR, Deamer DW (1978) Liposomes from ionic, single-chain amphiphiles. Biochemistry 17:3759–3768

Hargreaves WR, Mulvihill SJ, Deamer DW (1977) Synthesis of phospholipids and membranes in prebiotic conditions. Nature 266:78–80

Luisi PL (2003) Autopoiesis: a review and a reappraisal. Naturwissenschaften 90:49–59

Luisi PL (2006) The emergence of life: from chemical origins to synthetic biology. University Press, Cambridge

Mansy SS, Szostak JW (2008) Thermostability of model protocell membranes. Proc Natl Acad Sci USA 13351–13355

Mansy SS, Schrum JP, Krishnamurthy M, Tobe S, Treco DA, Szostak JW (2008) Template-directed synthesis of a genetic polymer in a model protocell. Nature 454:122–125
Miller SL (1953) A production of amino acids under possible primitive earth conditions. Science 117:528–529
Morowitz HJ (1992) Beginnings of cellular life. Metabolism recapitulates biogenesis. Yale University Press, New Haven
Oparin AI (2003) The origin of life. Dover, Mineola
Oro J (1961) Mechanism of synthesis of adenine from hydrogen cyanide under possible primitive earth conditions. Nature 191:1193–1194
Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242
Ricardo A, Carrigan MA, Olcott AN, Benner SA (2004) Borate minerals stabilize ribose. Science 303:196
Sacerdote MG, Szostak JW (2005) Semipermeable lipid bilayers exhibit diastereoselectivity favoring ribose. Proc Natl Acad Sci USA 102:6004–6008
Sun K, Mauzerall D (1996) A simple light-driven transmembrane proton pump. Proc Natl Acad Sci USA 93:10758–10762
Szathmary E (2007) Coevolution of metabolic networks and membranes: the scenario of progressive sequestration. Philos Trans R Soc Lond B Biol Sci 362:1781–1787
Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390
Walde P, Goto A, Monnard PA, Wessicken M, Luisi PL (1994a) Oparin's reactions revisited: enzymatic synthesis of poly(adenylic acid) in micelles and self-reproducing vesicles. J Am Chem Soc 116:7541–7547
Walde P, Wick R, Fresta M, Mangone A, Luisi PL (1994b) Autopoietic self-reproduction of fatty acid vesicles. J Am Chem Soc 116:11649–11654
Zepik HH, Blochliger E, Luisi PL (2001) A chemical model of homeostasis. Angew Chem Int Ed Engl 40:199–202
Zhu TF, Szostak JW (2009) Coupled growth and division of model protocell membranes. J Am Chem Soc 131:5705–5713

E-mail: mansy@science.unitn.it

The Autotrophic Origins Paradigm and Small-Molecule Organocatalysis

Eric Smith, Harold J. Morwitz

Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501; Krasnow Institute for Advance Study, George Mason University, Fairfax, VA 22030

Keywords Autotrophic Origins • Organocatalysis • SMILES • Computational Chemistry

Autotrophy and heterotrophy: disentangling the issues

All origin of life paradigms suppose at some level that cells have incorporated organic molecules that were once of abiotic origin. The distinction between primordial organosynthesis through high-energy processes quite different from those of biochemistry, and geochemical processes posited to be continuous with biochemistry, has become framed as a distinction between "heterotrophic" and "autotrophic" origin scenarios. The naming draws on an analogy with heterotrophy versus autotrophy of organisms, which is

convenient but potentially misleading. Core anabolism is universal at the ecosystem level, meaning that organism autotrophy or heterotrophy is an *ecological* distinction rather than a chemical one. To cast an analogy between the heterotrophic organism in its ecological environment, and *all* of emerging life in an environment of distinct organic chemistry, overemphasizes the organization of cellular hierarchy as a model for the organization of life, and underemphasizes the aspects of biochemistry not directly linked to hierarchical control.

If we suppose that modern life has preserved part of the inventory of primordial organic molecules—for whatever reasons of either their chemistry or commitments to higher-level molecular assemblies—then the major chemical distinction between autotrophic and heterotrophic paradigms concerns whether their mechanisms of synthesis were conserved or replaced, and whether the original molecular inventory was similar to the universal core today. By "mechanism" here we refer to the substrate-level architecture of the pathway and the elementary bond transitions; the extreme efficacy and selectivity of biological catalysts is a separate problem of emergence of higher-order structures, which either paradigm must address.

The autotrophy/heterotrophy divide does not involve significant disagreement about sources and forms of abiotically produced organics. All of the following sources are believed to have contributed to early-earth chemistry: low-temperature synthesis on dust and asteroid surfaces or cometary ices delivered by impacts (Kanavarioti et al. 2001); free radical and ionization reactions in the atmosphere (Miller 1953; Miller and Orgel 1974), and reducing reactions at the tectonically active lithosphere-hydrosphere interface (Russell and Martin 2004; Russel and Hall 2006; Martin and Russell 2006; Martin et al. 2008), potentially augmented with catalytically active mineral surfaces (Russell and Hall 1997). The problems for deciding between autotrophic and heterotrophic origins therefore come from uncertainty about mechanisms of *organization*. Which abiotic molecules and mechanisms could ever have become incorporated into networks capable of producing high molecular complexity and of permanently colonizing the geosphere? Did such incorporation depend on information encoded in higher-order structures such as oligomers, or was it more plausibly driven by chemical kinetics without hierarchical organization or explicitly informatic molecules?

The problem of the emergence, selection, and persistence of a biosynthetic network is not easily separated from questions about higher-order organization. The rate enhancement and selectivity of oligomer RNA and polypeptide catalysts are so powerful that—if their emergence from an unsupervised chemical medium were not so hard to explain—they might seem to provide a plausible route to replace practically any abiotic synthetic mechanism or to innovate any new pathway. The organization of such a "top-down" controlled metabolism would then most naturally be explained by a process of Darwinian selection. The assumption that this is the dominant organizing principle is implicit in many RNA-first scenarios for the origin of life, and has led to the explicit proposal that metabolism is a "palimpsest" of the RNA world (Benner et al. 1989). In our other abstract (on confederacy) we have mentioned probabilistic arguments (Simon 1973) that emergence and stability may be more likely for a cluster of partly autonomous metabolic modules than through a mechanism that relies heavily on top-down control.

Here we argue that a key quantitative feature of metabolism to be explained is the number of universal small core metabolites. It is about 300 (Srinivasan and Morowitz 2009)—a number much larger than the number of comparably complex molecules appearing with non-vanishing probability in a Gibbs equilibrium ensemble, but much smaller than the 10^7 molecules of

comparable complexity indexed by PubChem [http://pubchem.ncbi.nlm.nih.gov/]. Whatever organizational mechanism led from abiotic organosynthesis to biochemistry must have enabled significant complexity, while severely limiting the elaboration of that complexity. A restatement of the observation is that metabolism is *sparse* within the chemical possibility space, and that this sparseness may be seen as a result of *selection*. The key question then becomes what part of this selection was performed by chemical kinetics prior to hierarchy or acting independently of it, and what part was performed by Darwinian mechanisms using hierarchical integration and top-down control. Whether an organizing mechanism preserves or replaces molecules and synthetic pathways will determine whether our understanding of extant biochemistry is relevant to reconstructing stages of origination.

We currently lack a principled, quantitative, chemically explicit theory of the requirements to stabilize a metabolic network and a self-generated control system far from thermodynamic equilibrium. However, the last 10 years has seen rapid growth in knowledge and understanding of small-molecule organocatalysis (MacMillan 2008; Barbas 2008), which may offer new inputs to a theory of non-hierarchical chemical self-organization and selection. The term "organocatalysis" has been adopted to refer to catalysis by small organic molecules not making use of transition metals in the catalytic mechanism, as distinct from organo-metallic catalysis which is now an established field. It is not intended to include the study of macromolecular enzymes either, as this study is established within biochemistry.

Organocatalysis: a bridge from self-organization to evolution?

In autotrophic scenarios that credit a significant part of the organization of metabolism to primordially selected networks, the mechanism believed to account for the emergence and selection of reactions is autocatalysis. This may take the form of catalysis of reactions within the network by single molecules produced somewhere in the same network, or by topological motifs capable of amplifying their substrates (a phenomenon termed "network autocatalysis"). The idea is not new. It is the basis of Eigen's concept of the hypercycle (Eigen and Schuster 1977; Eigen and Schuster 1978), and has been applied in conceptual models by numerous researchers (de Duve 1991; Kauffman 1993; Segre et al. 1998a, Segre et al. 1998b; Segre and Lancet 1999; Segre et al. 2000; Segre et al. 2001). It is usually invoked for polymers, because their catalytic potential is widely recognized, but it is also usually represented abstractly, because the space of catalytic mechanisms is complex, and the prediction of catalytic power and selectivity from structural features is not well developed. Recent advances in the study *and generalization* (Barbas 2008) of mechanisms of catalysis by small molecules have the potential to convert abstract models of autocatalysis into quantitatively realistic simulations that can be used to guide experiment. Early work with explicit applications to origin of life was the survey by Pizzarello and Weber of the amino acids for catalysis of aldol condensations (Pizzarello and Weber 2004; Weber and Pizzarello 2006). Of interest in those studies was not only catalytic efficacy, but enantioselectivity with the potential to propagate chirality through reaction networks. More recent work has shown not only that small biomolecules can catalyze reactions, but also that the mechanisms of catalysis can often be related to known enzymatic mechanisms (Barbas 2008). Much of the work in this field has been driven by industrial synthetic chemistry, a highly quantitative discipline which has made heavy use of metallic and organometallic catalysts, and is now seeking lower-cost, non-hazardous catalysts to control synthesis and in some cases enantioselectivity (MacMillan 2008).

The distinct roles which the same molecule can fill within a non-hierarchical network, as catalyst or as reactant, opens the possibility for more nuanced theories of early biogenesis than a pure opposition between autotrophy and heterotrophy. Arguments for autotrophic origins often turn on the steady availability and high concentration of geochemically generated organics relative to deposits from space infall or atmospheric reactions (Shapiro 2006). More generally, the Oparin-Haldane conjecture (Fry 2000) that exogenous molecules fed the first higher-order assemblies requires that pathways have been discovered to produce particular molecular seed species before "the soup is exhausted" (Shapiro 2007). While we find these arguments convincing for the reactants in a network, they are less restrictive for its catalysts. The most robust exogenous organics, such as polycyclic aromatic hydrocarbons, graphenes, etc., could contribute to organic gels when convected through hydrothermal systems, serving as differential diffusion barriers, adsorbants, dielectric contrasts to aid phase-transfer catalysis, or alignment sites for nucleobases or other planar molecules. Not being consumed by reactions, they would be under less pressure from exhaustion, and their functions could successively be replaced by endogenously generated molecules and eventually by the whole micro-environments of protocells.

The distinction between catalyst and reagent opens a quantitative question about the transition from gechemical self-organization to early evolution: are catalysts in a network more easily replaced than reagents as the network flux and molecular complexity increase? We have proposed this asymmetry as a path from proto-metabolism to the RNA world (Copley et al. 2007), as a continuation of our proposal for an autocatalytic loop between nucleotide biosynthesis (from amino acid precursors) and amino acid synthesis from citric-acid cycle precursors (with nucleotides acting as catalysts) (Copley et al. 2005). This proposition has the corollary that the fitness of oligomers would be dominated by their ability to support existing pathways rather than by competitive self-replication, causing complex life to preserve rather than to over-write primordial pathways. It will be important to test this proposition both for particular molecules and as a general principle of network evolution.

Autocatalytic small-molecule networks: data and models

Two research thrusts are immediately suggested, to develop our knowledge of organo-catalysis and to apply it to the origin of life. The first is a comprehensive survey of the small universal core metabolites (Srinivasan and Morowitz 1999) for catalytic efficacy, specificity, and enantioselectivity on standard organic reactions. Priority may be given to those reactions that make up the bulk of core biosynthesis and which are supported in extant life by cofactors as group-transfer agents. This survey could be coordinated by a small core of researchers to be performed in parallel across a wide range of teaching institutions looking for low-cost, finite scope projects. Perhaps the SETI@home sky survey [http://setiathome.ssl.berkeley.edu/] provides a model for task division and aggregation of results.

The second thrust is to develop the theory of large-network autocatalysis with realistic chemical stoichiometry. An important difference between origin-of-life questions and those that arise in industrial synthesis is that industrial organocatalysts—like their metallic precursors—are exogenously supplied and controlled, whereas positive feedbacks and non-equilibrium network growth are of primary interest to origins. Network simulation by origins researchers has so far depended heavily on numerical simulation or simplification of reaction stoichiometry, but these limitations may be surmountable. Algorithms currently exist to generate molecules in SMILES format from SMIRKS [http://www.daylight.com/] or other representations of reactions (Benko et al. 2005). Working from a basis of such networks and a database of catalytic species or functional groups, it may be possible to

estimate whether positive feedbacks are rare or common, and whether they have the effect of focusing flows of energy and materials so as to select sparse networks comparable to those in biochemistry. If the broad statistical features of large chemical networks can be characterized and empirically validated, monte carlo models and simulated dynamics may become useful to make quantitative predictions.

The common theme in our proposals is to increase the heterogeneity of model elements without sacrificing chemical realism, so that our model systems will become more plausible approximations to early earth. This theme is not limited to organocatalysis. Knowledge gained in biochemistry about mechanisms of transition-metal catalysis in organic molecules—particularly cooperative mechanisms involving atoms of two different metals—may be used to revisit catalysis by natural minerals. Surveys of mineral catalysis of biologically relevant reactions could provide both initial conditions and parallel inputs for networks organized by organocatalysts.

References

Barbas CF III (2008) Organocatalysis lost: modern chemistry, ancient chemistry, and an unseen biosynthetic apparatus. Angew Chemie Int Ed 47:42–47

Benko G, Flamm C, Stadler PF (2005) Explicit collision simulation of chemical reactions in a graph based artificial chemistry. European Conference on Artificial Life, pp 725–733

Benner SA, Ellington AD, Tauer A (1989) Modern metabolism as a palimpsest of the RNA world. Proc Nat Acad Sci USA 18:7054–7058

Copley SD, Smith E, Morowitz HJ (2005) A mechanism for the association of amino acids with their codons and the origin of the genetic code. Proc Nat Acad Sci USA 102:4442–4447

Copley SD, Smith E, Morowitz HJ (2007) The origin of the RNA world: co-evolution of genes and metabolism. Bioorganic Chemistry 35:430–443

de Duve C (1991) Blueprint for a cell. Patterson, Burlington

Eigen M, Schuster P (1977) The hypercycle, part a: the emergence of the hypercycle. Naturwissenschaften 64:541–565

Eigen M, Schuster P (1978) The hypercycle, Part C: The realistic hypercycle. Naturwissenschaften 65:341–369

Fry I (2000) The emergence of life on earth: a historical and scientific overview. Rutgers U, New Brunswick

Kanavarioti A, Monnard P-A, Deamer DW (2001) Eutectic phases in ice facilitate nonenzymatic nucleic acid synthesis. Astrobiology 1:271–281

Kauffman K (1993) The origins of order: self-organization and selection in evolution. Oxford, London

MacMillan DWC (2008) The advent and development of organocatalysis. Nature 455:304–308

Martin W, Russell MJ (2006) On the origin biochemistry at an alkaline hydrothermal vent. Phil Trans Roy Soc B doi:10.1098/rstb.2006.1881, pp 1–39

Martin W, Baross J, Kelley D, Russell MJ (2008) Hydrothermal vents and the origin of life. Nature Rev Microbiol 6:805–814

Miller SL (1953) Production of amino acids under possible primitive earth conditions. Science 117:528–529

Miller SL, Orgel LE (1974) The origins of life on earth. Prentice-Hall, Engelwood Cliffs

Pizzarello S, Weber AL (2004) Prebiotic amino acids as asymmetric catalysts. Science 303:1151

Russell MJ, Hall AJ (1997) The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and ph front. J Geol Soc London 154:377–402
Russell MJ, Martin W (2004) The rocky roots of the acetyl-coa pathway. Trends Biochem Sci 29:358–363
Russell MJ, Hall AJ (2006) The onset and early evolution of life. Geol Soc Am Memoir 198:1–32
Segre D, Lancet D (1999) A statistical chemistry approach to the origin of life. Chemtracts, Biochem Mol Biol 12:382–397
Segre D, Lancet D, Kedem O, Pilpel Y (1998a) Graded autocatalysis replication domain (gard): kinetic analysis of self-replication in mutually catalytic sets. Orig Life Evol Biosphere 28:501–514
Segre D, Pilpel Y, Lancet D (1998b) Mutual catalysis in sets of prebiotic organic molecules: evolution through computer simulated chemical kinetics. Physica A 249:558–564
Segre D, Ben-Ali D, Lancet D (2000) Compositional genomes: Prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Nat Acad Sci USA 97:4112–4117
Segre D, Shenhav B, Kafri R, Lancet D (2001) The molecular roots of compositional inheritance. J Theor Biol 213:481–491
Shapiro R (2006) Small molecule interactions were central to the origin of life. Q Rev Biol 81:105–125
Shapiro R (2007) A simpler origin for life. Sci Am, Feb 12 (online)
Simon HA (1973) The organization of complex systems. In: Pattee HH (ed) Hierarchy theory: the challenge of complex systems. George Braziller, New York, pp 3–27
Srinivasan V, Morowitz HJ (2009) The canonical network of autotrophic intermediary metabolism: Minimal metabolome of a reductive chemoautotroph. Biol Bulletin 216:126–130
Weber AL, Pizzarello S (2006) The peptide-catalyzed stereospecific synthesis of tetroses: a possible model for prebiotic molecular evolution. Proc Nat Acad Sci USA 103:12713–12717

E-mail: desmith@santafe.edu

On Question 3: Heterotrophic Versus Autotrophic Scenarios

Theodor Alpermann,[1] Kristin Rüdel,[2] Ronny Rüger,[2] Nikolaus Knoepffler,[3] Alfred Fahr[2], Wolfgang Weigand[1]

[1]Institut für Anorganische und Analytische Chemie, Friedrich-Schiller-Universität Jena, Germany; [2]Lehrstuhl für Pharmazeutische Technologie, Friedrich-Schiller-Universität Jena, Germany; [3]Ethikzentrum Jena, Lehrstuhl für Angewandte Ethik, Friedrich-Schiller-Universität Jena, Germany

Keywords Autotrophy • Heterotrophy • Iron-sulfur-world • Surface Metabolism • Compartmentalization

How first life managed to power its metabolism is still a controversially discussed issue. Did the first life resort on already abundantly available biochemical building blocks? Alternatively, did the first life tediously have to produce the required building blocks for its

self-sustainment? Whatever hypothesis claims to explain the origin of life, a plausible proposal for the formation, alteration and evolution of biomolecules starting from simple compounds like amino acids towards oligomeric and polymeric has to be proved further. Such considerations include plausible sources of the "bio" elements (carbon, nitrogen, phosphorus, sulfur etc.) as well as an energy source that was available and sufficiently specific to generate biomolecules and to keep a metabolism running.

The hypothesis of a heterotrophic origin of life on Earth is strongly supported by the famous *Miller* experiment in the early 1950s (Miller 1953). The *Miller* experiment was an indication for a possible heterotrophic origin of life on planet Earth because of its straightforward experimental setup as well as the surprisingly large variety of resulting biochemical building blocks. In the following decades, more or less similar primordial conditions were simulated in reactions showing the feasibility to produce almost any relevant extant biochemical building block. Nevertheless, several open questions concerning the plausibility of the prebiotic broth hypothesis remain. It is still not explained how simple building blocks (e.g. amino acids or nucleobases) can undergo condensation reactions in diluted aqueous solutions, as postulated by the prebiotic broth hypothesis, and how polymeric compounds were originated finally (Miller 1959). According to thermodynamics, hydrolysis reactions (from peptides to single amino acids) are favoured in aqueous environments. Moreover, even if significant amounts of organic matter (including tar) were formed under the conditions of *Miller*'s experiment, the concentration of dissolved simple organic precursors in a prebiotic broth would have been probably too low for further reactions e.g. polymerisation for self-replication. Besides, the early atmospheric conditions are subject to discussions. Recent research showed, that the primitive Earth's atmosphere did contain mainly CO_2, N_2 and a certain amount of H_2 (Tian 2005). Such a rather neutral atmosphere is completely different from the proposed reducing atmosphere of the broth hypothesis. Thus, central issues of the broth hypothesis lack a reasonable explanation.

In contrast, there are some important experiments that support the idea of a chemoautotrophic origin of life. Such a scenario was proposed by *Wächtershäuser*'s Iron-Sulfur-World hypothesis (Wächtershäuser 1988; Wächtershäuser 1992). The development of a surface metabolism at an iron sulphide surface, which is driven by the reaction (1) is assumed as a possible ancestor of the first life. At an iron sulphide surface, organic matter would have been reduced by the FeS / H_2S system and in situ bound to the positively charged surface of the formed pyrite. The properties of the so-called "pyrite-pulled" reaction system of FeS and H_2S were extensively investigated (Rickard 1997a; Rickard 1997b; Rickard 2007; Taylor 1979) and the reductive power was further demonstrated in several intriguing experiments.

$$FeS + H_2S \rightarrow FeS_2 + H_2; \quad \Delta_R G^\circ = -38\mathrm{KJ/mol},\ \mathrm{pH}\ 0 \tag{1}$$

$$CO + H_2O + CH_3SH \rightarrow CH_3COOH + H_2S; \quad \text{catalyst: (Fe, Ni)S} \tag{2}$$

$$N_2 + 3H_2S + 3FeS \rightarrow 2NH_3 + 3FeS_2; \Delta_R\, G^\circ < 0 \tag{3}$$

According to the hypothesis the reduction of CO_2 by the FeS / H_2S system results in surface bound formate. Experimentally, only the formation of methyl mercaptane and simple sulfur derivatives was observed (Heinen 1996). In subsequent experiments it was shown by

Wächtershäuser et al. that acetic acid can be formed from thioacetic S acid, using CH_3SH as methyl source and CO (**reaction 2**) in a catalytically promoted environment (Huber 1997). That reaction proceeds at 100°C in the presence of (Fe,Ni)S minerals. The reaction is similar to the acetyl-CoA biosynthetic pathway catalysed by Fe-S and (Fe,Ni)-S centers at the active part of the Acetyl-CoA synthase. Indeed, several important extant enzymes e.g. hydrogenases and CO-dehydrogenase contain Fe-S and (Fe,Ni)-S clusters. α-Amino acid activation and peptide formation were discovered when α-amino acids were reacted with CO and H_2S or CH_3SH under the same conditions (CO/(Fe,Ni)S/100°C) (Huber 1998). It has also been argued that the biochemical process of ammonia formation, from dinitrogen, by enzymes containing iron–sulfur clusters as the active centre may be traced back to a prebiotic "pyrite-pulled" nitrogen fixation. The possibility to reduce nitrogen (N_2) by the FeS / H_2S system (**reaction 3**) could be proved experimentally (Dörr 2003). The synthesis of ammonia may serve as a model for a primordial nitrogen fixing system and it conforms well to theories of a chemoautotrophic origin of life. Consequently, an important key experiment supporting the chemoautotrophic theory would be the successful reaction of CO_2 with H_2S in the presence of FeS and NiS yielding thioacetic-S-ester as well as a joint activation of CO_2 and N_2 forming amino acids. Hence, there is still a strong demand for further investigations of the reducing power of the FeS / H_2S system.

Compartmentalization is also proposed to be a major milestone for the development of life (Oberholzer 1999). To investigate both the process of compartmentalization and the iron sulfur world hypothesis, the development of a primordial cell unit containing the FeS / H_2S system would be promising.

There are even other hypotheses that deal with the crucial question of what came first, heterotrophy or autotrophy. Overall, there is little doubt that our knowledge of the primitive Earth around 3.8 billion years ago is very fragmentary.

Does this impossibility of falsification of any hypothesis mean that there is a philosophical hypothesis left, the hypothesis of a supernatural Creator, who brings the process of origin of life into being? This hypothesis is not falsifiable, either. But the difference to the heterotrophic and autotrophic or other scientific hypothesis is a fundamental one. Every scientific hypothesis can only be judged concerning its plausibility and, with certain restrictions, its experimental verifiability. The hypothesis of a supernatural Creator is not a scientific hypothesis. There is no possibility of an experimental verification. If it is used as if there were this kind of possibility, then a categorical failure is committed. Thus the origin of life may either resort on already abundantly available biochemical building blocks (heterotrophic origin of life) or the first life may tediously have to produce the required building blocks for its self-sustainment (chemoautotrophy).

References

Dörr M, Kässbohrer J, Grunert R, Kreisel G, Brand WA, Werner RA, Geilmann H, Apfel C, Robl C, Weigand W (2003) A possible prebiotic formation of ammonia from dinitrogen on iron sulfide surfaces. Angew Chem Int Ed Engl 13:1540–1543

Heinen W, Lauwers AM (1996) Organic sulfur compounds resulting from the interaction of iron sulfide, hydrogen sulfide and carbon dioxide in an anaerobic aqueous environment. Orig Life Evol Biosph 2:131–150

Huber C, Wachtershauser G (1997) Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. Science 5310:245–247

Huber C, Wachtershauser G (1998) Peptides by activation of amino acids with CO on (Ni, Fe)S surfaces: implications for the origin of life. Science 5377:670–672

Miller SL (1953) A production of amino acids under possible primitive earth conditions. Science 3046:528–529
Miller SL, Urey HC (1959) Origin of Life. Science 3389:1622–1624
Oberholzer T, Nierhaus KH, Luisi PL (1999) Protein expression in liposomes. Biochem Biophys Res Commun 2:238–241
Rickard D (1997a) Kinetics of pyrite formation by the H2S oxidation of iron (II) monosulfide in aqueous solutions between 25 and 125 degrees C: the rate equation. Geochim Cosmochim Acta 1:115–134
Rickard D, Luther GW (1997b) Kinetics of pyrite formation by the H2S oxidation of iron (II) monosulfide in aqueous solutions between 25 and 125 degrees C: the mechanism. Geochim Cosmochim Acta 1:135–147
Rickard D, Luther GW (2007) Chemistry of iron sulfides. Chem Rev 2:514–562
Taylor P, Rummery TE, Owen DG (1979) Reactions of iron monosulfide solids with aqueous hydrogen sulfide up to 160°C. J Inorg Nucl Chem 12:1683–1687
Tian F, Toon OB, Pavlov AA, De SH (2005) A hydrogen-rich early Earth atmosphere. Science 5724:1014–1017
Wächtershäuser G (1988) Before enzymes and templates: theory of surface metabolism. Microbiol Mol Biol Rev 4:452–484
Wächtershäuser G (1992) Groundworks for an evolutionary biochemistry: the iron-sulphur world. Prog Biophys Mol Biol 2:85–201

E-mail: wolfgang.weigand@uni-jena.de

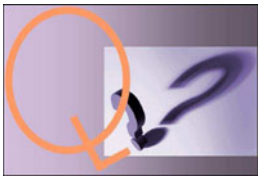
Workshop OQOL'09

Extended Abstracts for the Following Selected Question

- **On the Origin of Catalytic Cycles**

Premise. In a prebiotic scenario, like that assumed by Stanley Miller in his famous experiments, once given the initial conditions, prebiotic reactions flow towards the most stable compounds, being ruled by thermodynamic control. With the 'free ticket' of thermodynamic control, however, chemical prebiotic evolution would not have gone very far. In fact, the question of the origin of life can be abstracted as the question of the origin of enzyme-like controlled catalysis (eventually leading to genetically controlled catalysis), giving rise to sequential metabolic cycles, as opposite to chemically equilibrated reaction pathways. One line of thought considers that films of organic materials, found bound to the hot internal surfaces of inorganic tubes in contemporary hydrothermal systems, may have initiated networks of interaction between different layers that led the way towards metabolic cellular life.

The question. How do you envisage the origin of sequentially catalytized reactions in a prebiotic scenario? And can you provide facts or scientific arguments, not simply beliefs, about this critical point?



On the Origin of Catalytic Cycles

Stuart Kauffman

Institute for Biocomplexity and Informatics, University of Calgary, Canada

Keywords DNA Collectively Autocatalytic Sets. Peptide Collectively Autocatalytic Sets • Laws of Organization • Emergence of Collectively Autocatalytic Sets • Maximizing Power Efficiency • Dynamical Criticality

I begin by noting that catalytic cycles are an accomplished experimental fact. Guenter von Kiedrowski has made collectively autocatalytic sets of single stranded DNA molecules. Reza Ghadiri has made collectively autocatalytic sets of peptides. In these collectively autocatalytic sets, a molecule, A, catalyses the formation of B from B fragments, and B catalyzes the formation of A from A fragments. Ghadirii's results establish firmly that molecular reproduction need not be based on template replication of DNA, RNA or similar molecules although such replication without enzymes is not ruled out in the origin of life. The next issue is the range of molecules that might play the roles of catalysts in catalytic cycles. These include RNA, peptides, and organometalic molecules such as peptides with metalic moieties. The probability that a "randomly chosen" molecule in these diverse classes can function as a catalysis may differ dramatically. Experiments with in vitro evolution of RNA ribozymes suggests a probability that a random RNA sequence catalyzes a specific reaction at about one in a thousand trillion. For random peptides, due to the higher chemical diversity, and data showing that random peptides bind to arbitrary epitopes with a probability of about one in a million, I would hazard the guess that the probability a random peptide catalyzes a randomly chosen reaction at about one in a billion. Julies Rebek guesses that the chances an organometalic compound catalyzes a random reaction may be about one in a thousand. Tetsuya Yomo at this conference, informed me that he has shown that random peptides length about 143 have a 1% chance of catalyzing a reaction with low efficiency. We are working together now to make collectively autocatalytic peptide sets, as described further below.

My own work has uncovered what may be a law of organization. Consider a set of molecules that can serve as substrates and products of reactions, and are themselves candidates to catalyze those very reactions. The ratio of molecular diversity to reactions among a set of N different species depends upon the order of the reaction. For cleavage and ligation reactions among linear polymers the ratio of reactions to molecules scales as the length of the longest polymer, say P. For two substrate two product reactions, and if any pair of molecules can undergo at least one two substrate two product reaction, the number of reactions scales as the square of the diversity of molecular species, hence the ratio of reactions to molecules scales as the diversity of molecular species, say N. N rises much faster than P as linear polymer length and diversity of possible polymers increase. The ratio is unknown for non-linear polymers. Numerical and analytic work show that if the probability of catalysis of a given reaction by a given molecule is above a threshold, then as the diversity of molecular species increases, more and more reactions are catalyzed and a "giant connected catalyzed reaction system" arises. With probability approaching 1.0, this giant component will contain a collectively autocatalytic set. Obviously the size of this set depends critically on the probability that molecules catalyze reactions, so is far smaller and simpler to obtain if the probability of catalysis is high. This suggests organometalic peptides as candidates of choice. But Yomo's results also argue that random polypeptides may work well too.

Once catalysis is achieved, it can also be inhibited, either competitively or non-competitively. This unstudied problem will yield autocatalytic sets with complex dynamics. Work on random causal (Boolean) systems shows that these can behave in three regimes, ordered, critical, and chaotic. It may be deeply important that critical networks store information, propagate information, and correlate the most complex and diverse behavior among variables whose activities or concentrations can vary. If dynamically critical autocatalytic sets can exist—and there is evidence cells are dynamically critical—they may be able to coordinate the most complex behaviors within and between such co-evolving sets.

Another major issue is energy flow through such autocatalytic sets. I have elsewhere suggested that real cells do work cycles and defined a "molecular autonomous agent" as an open thermodynamic system able to do reproduce and do at least one work cycle. Now work cycles maximize energy efficiency, as Carnot showed, if done adiabatically, ie infinitely slowly. But such systems, if the work cycle is needed for reproduction, would lose the Darwinian kinetic race that Addy Pross talked about at this conference. Thus, energy efficiency is not the right concept. In its place I hypothesize that cells and early life maximizes something like a power efficiency per metabolic fuel consumed. For an automobile, miles per gallon is maximized, not at 2 or 2,000 miles per hour, but at say 47 miles per hour. This is deeply interesting because it picks out a finite displacement from equilibrium, beyond the reversible Onsager relations near equilibrium, and beyond Prigogine's bifurcation a short distance from equilibrium. Now the analogue of miles per gallon is something like cell reproduction rate divided by metabolic rate. Yomo and I are now testing for a maximum of this ratio, plotted on the Y axis, versus reproduction rate on the X axis. We hope that entropy production, measured as heat production, is minimized where E. coli maximizes reproduction rate/metabolic rate, and that quarum sensing holds bacterial colonies at this point. If so, then there may be a hint of a general law: it costs energy to obtain energy, so cells and organisms may tend to maximize their reproduction rate given metabolic costs.

A major final point is that the emergence of collectively autocatalytic sets, now testable using "never before born proteins" or organometalic peptides, is what Robert Laughlin calls a "Law of Organization", not reducible to fundamental physics, but a law in its own right. As such, it is emergent with respect to fundamental physics, as I will describe in my answer to question 2.

E-mail: skauff@telus.net

Properties of Catalytic Cycles

Peter Schuster

Theoretical Biochemistry Group, University of Vienna, Austria

Keywords Autocatalytic Cycles • Cyclic Catalysis • Genetic Regulatory Networks • Hypercycles • Metabolic Cycles • Predator-prey Systems

Feedback loops are among the simplest known systems with properties that are not shown by its elements: (i) A cycle of uncatalyzed reactions acts as a catalyst. (ii) A cyclic network of reactions catalyzed by members of the cycle represents a composite

autocatalyst. (iii) A cycle of autocatalytic reactions catalyzed by the autocatalysts shows new properties like hyperbolic growth and integration of otherwise competing elements (Eigen & Schuster 1979).

Cycles of reactions are often discussed as a possibility for reducing the combinatorial manifold of organic molecules to the building blocks of present day life (Schuster 2000). It is intriguing to assume suitable reaction sets that convert a substrate into a product whereby the set of intermediates in total remains unchanged. A seemingly useful example is the reverse or reductive citric acid cycle with the stoichiometry

$$\text{citrate}^{3-} + 6\ CO_2 + 9\ H^+ \rightarrow 2\ \text{citrate}^{3-} + 5\ H_2O + 3\ H^+$$

Since citrate is member and product of the catalytic cycle, the reductive citric acid cycle is autocatalytic and would be ideal for canalizing a plethora of possible alternative reaction products into citrate, a key compound in the synthesis of several amino acids and other important metabolites. Indeed, the reductive citric acid cycle (Morowitz et al. 2000; Smith & Morowitz 2004) and variants of it involving sulfur-containing compounds (Wächtershäuser 1990) have been suggested as solutions to the early metabolism problem but so far not a single successful enzyme-free experimental implementation was reported. In contrast the enzyme-catalyzed reductive citric acid cycle is abundant in microorganisms living at deep-sea hydrothermal vents (Campbell & Cary 2004). A recent photochemical approach involving ZnS particles was successful for individual reaction steps but unsuccessful for the cycle as a whole (Zhang & Martin 2006). In a critical account on the plausibility of the reaction cycles proposed in prebiotic chemistry, Leslie Orgel (2000), states that the only self-organizing biochemical cycle operating experimentally is the formose reaction which, however, has many side reactions and gives rise to a great variety of different carbohydrate compounds.

Autocatalytic sets of proteins, in which all members are—at the same time—substrates and catalysts have been proposed 40 years ago by Stuart Kauffman (1969; 1986). No experimental example supporting the occurrence of large cycles has been reported so far. Only single peptide replicators (Lee et al. 2006) and cycles with two members showing cross catalysis were designed and synthesized (Lee et al. 2007). In case of enzyme-free replication of oligonucleotides the first molecule found to undergo self-replication was a deoxyribo-hexanucleotide assembled from two trinucleotide precursors (von Kiedrowski 1986). A cycle of two DNA oligomers with cross-catalysis has been successfully designed as well (Sievers & von Kiedrowski 1994). Further breakthrough was achieved in 2009 through the design and experimental implementation of a cross catalytic cycle in which two ribozymes are synthesized from four precursor RNA molecules (Lincoln & Joyce 2009). Populations of cross-replicating RNA enzymes show mutation, recombination, and adaptation to niches (Voytek & Joyce 2009). This system is a highly interesting candidate for evolution in a pure RNA world.

It is worth mentioning that attempts to design cyclic genetic control networks with more than two members have been successful. In such networks (regulatory) proteins activate or inhibit the synthesis of other (regulatory) proteins in the sense of protein sets discussed above. As an example we mention the repressilator, a system with cyclic inhibition of three genes, which was designed and expressed in *E. coli* (Elowitz & Leibler 2000) and which shows the expected dynamic behavior, e.g. oscillations (Müller et al. 2006).

Small cycles of interacting autocatalysts are common in ecosystem biology. Examples with two members fall into two classes: (i) hypercycles or symbioses and (ii) predator-prey systems. In its elementary form a hypercycle with two members is a system in which the replication of a species is dependent on the presence of the second species. This implies that the replication term in the differential equations is proportional to the concentrations of both species:

$$\frac{dx_1}{dt} = x_1(g_{12}x_2 + \cdots) \quad \text{and} \quad \frac{dx_2}{dt} = x_2(g_{21}x_1 + \cdots) \quad \text{with} \quad g_{12}, g_{21} > 0$$

Many examples of these 'hard' symbioses are known in nature. So far no molecular implementations of such systems were reported. Predator-prey systems have been modeled first by Alfred Lotka (1920) and Vito Volterra (1926). They are characterized by the same form of interaction terms but different signs: $g_{12}>0$ and $g_{21}<0$. In other words, presence of prey is essential for predator replication and presence of predator has a negative effect on the abundance of prey. An attempt to implement a predator-prey system by means of nucleic acid molecules has been successful (Wlotzka & McCaskill 1996; Ackermann et al. 1998). Practically no data are available from experiments with molecules for well understood systems with more than two elements.

The final questions to be answered are: Why do cyclic chemical reaction networks only rarely or not at all work without enzymes and how do proteins facilitate the operation of catalytic and autocatalytic cycles? Two general difficulties seem to be hard to circumvent without enzymes: (i) High reaction specificity is required for avoiding diversification on a wide network of reactions, and (ii) completion of a cycle requires vanishing Gibbs free energy and any downhill cascade of reactions from substrate to product has to be compensated by uphill reactions requiring tuning of rate parameters in order to avoid too slow reaction steps. The first point can be addressed by means of an example from astrophysics, the Bethe-Weizsäcker or CNO-cycle that accounts for part of the helium production (1.7%) in the sun (Bethe 1939):

$$4\ {}^{1}\mathrm{H} \rightarrow ({}^{12}\mathrm{C}, {}^{13}\mathrm{N}, {}^{13}\mathrm{C}, {}^{14}\mathrm{N}, {}^{15}\mathrm{O}, {}^{15}\mathrm{N}, {}^{12}\mathrm{C}) \rightarrow {}^{4}\mathrm{He} + 2\mathrm{e}^{+} + 2\mu_{\mathrm{e}} + \gamma + 26.8\ \mathrm{MeV}$$

The six nuclei in parentheses form the core of intermediates in the catalytic cycle. Nuclear reactions are highly specific but despite specificity the CNO-cycle has two competitors with the cores (^{15}N,^{16}O,^{17}F,^{17}O,^{14}N,^{15}O,^{15}N), having a frequency of 0.04% in the sun relative to the CNO-cycle, and (^{17}O,^{18}F,^{18}O,^{19}F,^{16}O,^{17}F,^{17}O) being significant only in massive stars. The difference to chemical reactions, however, becomes obvious in the overall reaction, which is identical in all three cases. This is not true for organic chemistry where branching reactions give rise to an increasing diversity of products that weakens the cycles. Specificity is introduced by protein catalysts, which block all undesired side reactions. The second point is also solved by catalysis: High activation energies may slow down some reactions, in particular uphill reactions, such that the cycle can't be completed efficiently. Both problems do not hit insurmountable physical barriers and should be solvable, in principle, by means of non-protein catalysts and/or photochemical reactions too, but high intuition and

ingenious experimental skill will be required for finding systems operating under plausible prebiotic scenarios.

References

Ackermann J, Wlotzka B, McCaskill JS (1998) In vitro DNA-based predator-prey system with oscillatory kinetics. Bull Math Biol 68:329–353

Bethe, HA (1939) Energy production in stars. Phys Rev 55:434–456

Campbell BJ, Cary SC (2004) Abundance of reverse tricarboxylic acid cycle genes in free-living microorganisms at deep-sea hydrothermal vents. Appl Environ Microbiol 70:6282–6289

Eigen M, Schuster P (1979) The hypercycle. A principle of natural self-organization. Springer, Berlin

Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403:335–338

Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. J Theoret Biol 22:437–467

Kauffman SA (1986) Autocatalytic sets of proteins. J Theoret Biol 119:1–24

von Kiedrowski G (1986) A self-replicating hexa deoxynucleotide. Angew Chem Int Ed Eng 25:932–935

Lee DH, Granja JR, Martinez JA, Severin K, Ghadiri MR (1996) A self-replicating peptide. Nature 383:525–528

Lee DH, Severin K, Yokobayashi Y, Ghadiri MR (1997) Emergence of symbiosis in peptide self-replication through a hypercyclic network. Nature 390:591–594

Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. Science 323:1229–1232

Lotka AJ (1920) Undamped oscillations derived from the law of mass action. J Am Chem Soc 42:1595–1599

Morowitz HJ, Kostelnik JD, Yang J, Cody GD (2000) The origin of intermediary metabolism. Proc Natl Acad Sci USA 97:7704–7708

Müller S, Hofbauer J, Endler L, Flamm C, Widder S, Schuster P (2006) A generalized model of the repressilator. J Math Biol 53:905–937

Orgel LE (2000) Self-organizing biochemical cycles. Proc Natl Acad Sci USA 97:12503–12507

Schuster P (2000) Taming combinatorial explosion. Proc.Natl.Acad.Sci.USA 97:7678–7680

Sievers D, von Kiedrowski G (1994) Self-replication of complementary nucleotide-based oligomers. Nature 369:221–224

Smith E, Morowitz HJ (2004) Universality in intermediary metabolism. Proc Natl Acad Sci USA 101:13168–13173

Volterra V (1926) Fluctuations in the abundance of a species considered mathematically. Nature 118:558–560

Voytek SB, Joyce GF (2009) Niche partitioning in the coevolution of two distinct RNA enzymes. Proc Natl Acad Sci USA 106:7780–7785

Wächtershäuser G (1990) Evolution of the first metabolic cycles. Proc Natl Acad Sci USA 87:200–204

Wlotzka B, McCaskill JS (1996) A molecular predator and its prey: coupled isothermal amplification of nucleic acids. Chem Biol 4:25–33

Zhang XV, Martin ST (2006) Driving parts of Krebs cycle in reverse through mineral photochemistry. J Am Chem Soc 128:16032–16033

E-mail: pks@tbi.univie.ac.at

On the Origin of Specific Macromolecular Sequences, Catalytic Cycles and the RNA World: The 'Ligand-Imprinting Hypothesis'

Takuya Ueda

Department of Medical Genome Sciences, University of Tokyo, Japan

Keywords Amino Acid Condensation • Enzyme • Ligand • Genetic Code

We have no clue to the solution on the amino acid condensation process in the primordial soup. Especially how proteins with ligand binding properties emerged has been totally unsolved. To address this issue I present the 'ligand-imprinting hypothesis', in which activated amino aids were concentrated in the vicinity of ligand and condensation reaction efficiently proceeded to generate polypeptide with binding activity to a substrate.

Two problems in the origin of life

In life, nucleic acid capable of storing the genetic information inseparably coordinate with protein in charge of biotical function. Thus, the origin of life is deemed to be the emergence of the linkage of these two molecules. Nucleic acid and protein are, of course, polymers of nucleotides and amino aids, respectively. If the condensation process of these monomers occurred simultaneously in the vicinity, linkage of genetic and functional molecules might be generated. How protein obtained molecular recognition capability during amino acid condensation process in prebiotic era remains as an enigma. The "ligand-imprinting hypothesis" opens a way to give an answer to these two problems of origin of life.

Activation of amino acid

In this model amino acids and nucleotide were first activated to higher energy level. We can find two candidates of activated amino acids in the extant translation machinery; one is amino acids-AMP and the other is aminioacyl-tRNA (see: Fig. 1a, b, c). Amino acid-AMP might be excluded due to the lack of nucleotide variation (genetic) in this molecule. Although amino acid-AMP is at higher energy level than aminoacyl-tRNA, the lack of capacity for storing the amino acid sequence information is a fatal flaw of the molecule. In contrast, aminoacyl-tRNA is capable of encoding genetic information in RNA moiety. Of course, primitive tRNAs might be much shorter and simpler than present tRNA. It is likely that ancestral aminoacyl-tRNAs possessing various short ologonucleotide were plausible molecules for activated amino acids. The attached oligonuleotide might contain anticodon sequence, as shown in Fig. 1d. Assuming that anticodon portion of oligonulceotide has affinity toward corresponding amino acid, the activation may be accelerated by incorporating a cognate amino acid. Indeed, the anticodons in which hydrophobic adenosine is localized at the second letter, prefers hydrophobic aminio acids in the genetic code, suggesting a chemical relationship between anticodon and amino acid. Considering all the tRNAs have CCA sequences at the 3′-end without exception, it is possible to speculate that CCA played the role of a ribozyme-attaching-amino acid.

Ligand-imprinting hypothesis

In the second step, these activated amino acids with oligonucleotide disposed around a particular ligand molecule prior to polymerization reaction. By a condensation reaction of activated amino acids, ligand-embedded polypeptides were generated as molecules with binding properties (see Fig. 2). The ligand was dissociated from polypeptide leaving a biding pocket consisting of several amino residues. If the condensation of oligonucleotides took place coupling with amino acid condensation, the resultant polynucleotide corresponded to an amino acid sequence and can store primitive genetic information of polypeptide.

Conclusion

The ligand–imprinting provides several advantages in explaining origin of life and, in particular, the transition from an RNA-world to a protein-world. First, a primitive life needs not test all the possible amino acid sequences to obtain functional proteins. Second, the coupling of ligation of oligonucleotide and amino acid condensation makes the linkage between genotype of nucleic acids and phenotype of proteins. Third, the affinity of adapter RNA to amino acid might create a correspondence of anticodons to amino acids, and the establishment of the genetic code.

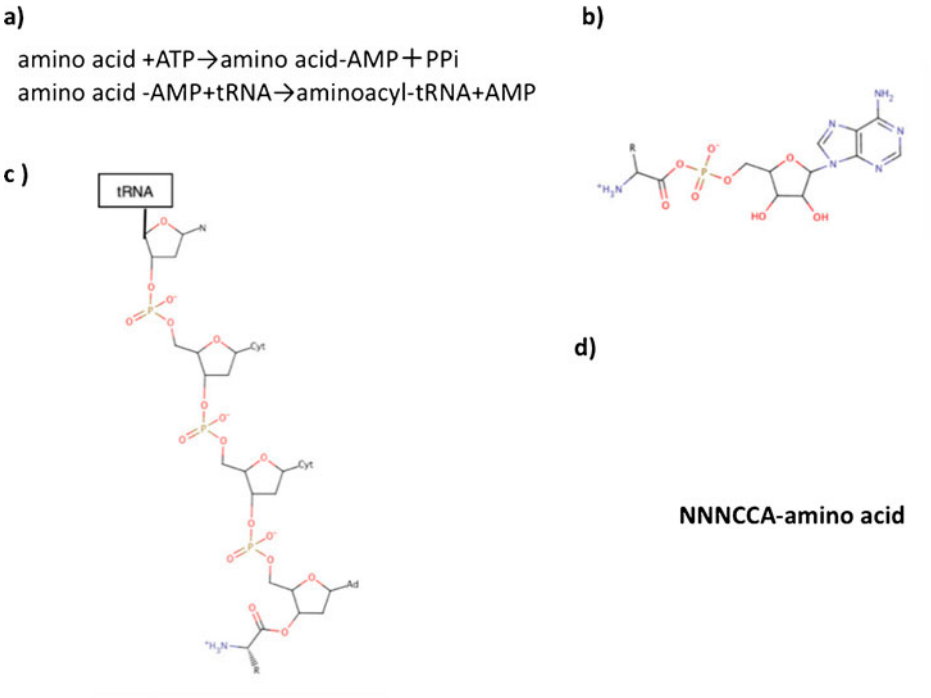


Fig. 1 Activation of amino acids a) Aminoacylation reaction b) The structure of amino acid-AMP c) The structure of extant aminoacyl-tRNA d) Putative primitive aminoacyl-tRNA NNN represents anticodon sequence

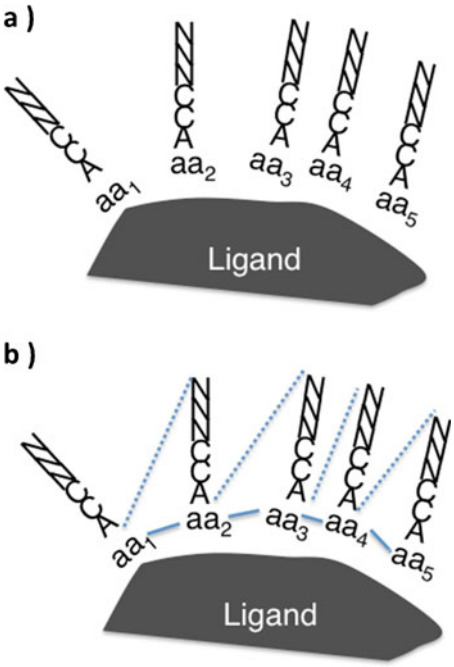


Fig. 2 Ligand-imprinting process of primitive protein a) Diposition of activated amino acids in the vicinity of ligand b) Simultaneous occurrence of amino acid condensation and oligonucleotide ligation. The broken and solid lines represent phosphodiester bond and peptide bond, respectively

E-mail: ueda@molbio.t.u-tokyo.ac.jp

On the Origin of Catalytic Cycles

Robert Shapiro

Department of Chemistry New York University, USA

Keywords Metabolism First • Reductive Citric Acid Cycle • Driver Reaction • Redox Couple • Serial Transfer

"Metabolism-first" theories of the origin of life propose that a self-reproducing collection of small molecules preceeded replicating polymers in the development of life. This concept has often been described in terms of an autocatalytic reaction cycle, in which sufficient quantities of carbon dioxide or of other simple organic molecules are absorbed in each turn of the cycle to double the amount of material within it (Kauffman 1994). The participating members of the cycle also serve as catalysts for the reactions of the cycle. Variants of the

reductive citric acid cycle have often been cited as possible examples of such a system (Wächtershäuser 1990; Morowitz 1999).

The plausibility of such a system has been challenged on a number of grounds. (Pross 2004; Orgel 2008). Many alternative possibilities for chemical reaction undoubtedly existed for organic molecules on the early Earth. Many of them would serve to drain material from the cycle, rather than sustain it. Any catalysts that were present would be as likely to facilitate these side reactions as they would the core reactions of the cycle. In the absence of specific enzymes, the organic material present would be likely to form a host of different substances and polymeric tars, as has been found in meteorites.

These objections can be remedied if an external energy source can be coupled specifically to a reaction of the central cycle (Morowitz 1968; Feinberg and Shapiro 1980). Thermodynamic factors would then favor the central cycle and draw organic material from competing reactions into it; no specific catalysis would be required. Environmental changes could lead to the evolution of the central cycle into a more complex self-sustaining reaction network. At some stage, the segregation of the reactive components within a suitable compartment would provide the first primitive cell. (Shapiro 2006).

While these ideas may be plausible, a "proof of principle" experiment will be needed to validate them. Some advocates of autocatalytic cycles have attempted to specify the participating components in advance, but this approach has not as yet proved fruitful. I feel that a more empirical approach should be used. An initial mixture of simple reactive organic molecules should be combined with an energy source (such as a redox couple), and the changes in composition of the mixture over time should be followed.

Initially, the mixture would become more chaotic as all of the available reaction paths were explored. In an unsuccessful run, all of the organic material would eventually become converted into unreactive, insoluble or unavailable materials (alkanes, precipitates, tars, escaping gases). Despite the continued supply of energy, no further significant changes would be observed. If however, the mixture produced a reaction which specifically utilized the energy, and the products of this "driver" reaction formed a cycle which regenerated the substrate(s) for the productive reaction, a different course might be followed (Shapiro 2006). Material that was linked by equilibria to the components of the cycle would be drawn into the cycle. The mixture would be partitioned into the cycle components and some unavoidable waste material. If additional material and energy were input, the evolution of the initial cycle into a more efficient energy-utilizing network might be observed.

Ideally, this experiment should be run under conditions in which fresh supplies of organics and energy were input continually, and intractible wastes were removed. Flow reactors could be devised for this purpose, but as a start, a variation of the serial transfer procedure used by S. Spiegelman and his colleagues to follow RNA evolution could be used (Spiegelman et al. 1965). In the present case, the initial mixture would contain small reactive organic molecules such as aldehydes, amino acids and α-keto acids, and a redox couple such as sulfide-disulfide. After a preset time, an aliquot of the mixture would be withdrawn for analysis, and another aliquot (excluding tars and precipitates) would be added to a new reaction vessel containing another sample of the initial mixture. After several such serial transfers, the analysis should reveal which path was being followed. If a self-sustaining cycle emerged, then conditions might be varied in subsequent transfers to stimulate further evolution of the system.

Such experiments could not duplicate the events involved in the origin of life on Earth, which are lost to history. Further, some problems involved in the origin-of-life, such as the mechanism by which compartments were formed to limit loss of material by diffusion, would not be addressed here. But experiments of this type would connect the process of

molecular self-organization (sometimes called "chemical evolution") to the existing laws of chemistry and physics, both on Earth and elsewhere in the universe.

References

Feinberg G, Shapiro R (1980) Life Beyond Earth. Morrow, New York
Kauffman S (1994) At Home in the Universe. Oxford University Press, New York
Morowitz HJ (1968) Energy Flow in Biology. Academic Press, New York
Morowitz HJ (1999) A theory of biochemical organization, metabolic pathways, and evolution. Complexity 4:39–53
Orgel LE (2008) The Implausibility of metabolic cycles on the Prebiotic Earth. PloS Biol 6: 5–13
Pross A (2004) Causation and the origin of life: metabolism or replication first? Orig Life Evol Biosph 34:307–321
Shapiro R (2000) A replicator was not involved in the origin of life. IUBMB Life 49:173–176
Shapiro R (2006) Small molecule interactions were central to the origin of life. Quart Rev Biol 81:105–125
Spiegelman S, Haruna I, Holland, IB, Beaudreau G, Mills D (1965) The synthesis of a self-propagating and infectious nucleic Acid with a purified enzyme. Proc Natl Acad Sci USA 543:919–927
Wächtershäuser G (1990) Evolution of the first metabolic cycles. Proc Natl Acad Sci USA 87:200–204

E-mail: rs2@nyu.edu

On the Development Towards the Modern World: A Plausible Role of Uncoded Peptides in the RNA World

Matthew J. Belousoff, Chen Davidovich, Anat Bashan, Ada Yonath

Dept of Structural Biology, Weizmann Institute, Israel

Keywords Proto-ribosome • Peptide Bond Formation • Uncoded Peptides • Genetic-code Evolution

Arguably one of the most outstanding problems in understanding the progress of early life is the transition from the RNA world to the modern protein based world. One of the main requirements of this transition is the emergence of mechanism to produce functionally meaningful peptides and later, proteins. What could have served as the driving force for the production of peptides and what would have been their properties and purpose in the RNA world? The answer may seem immediately clear; proteins are better enzymes than ribozymes. However, modern proteins are only useful in their folded stated, whereas peptides need to reach a critical size and specific amino acid sequence before they fold into a functional biomolecules. It is more than likely that emerging peptides were not immediately useful as operational enzymes. Herein we describe two plausible roles that emerging peptides could have played, firstly support of the pre-existing RNA machinery and secondly as early chemical catalysts.

During the era of the RNA world, protein (or peptide) evolution would have been strongly coupled to the extant RNA infrastructure development. Thus, the emergence of peptides as a mechanism for supporting pre-existing RNA machinery, is a sound reason for peptides to be retained in an RNA world. In this issue we briefly discuss ideas for the apparatus that would have served as the proto-ribosome (See Bashan et al. this volume), hypothesized to be a symmetrical 'pocket-like' RNA dimer capable of simple peptidyl

transfer and elongation, in essence non-coded peptide production (Agmon, et al. 2006, Agmon et al. 2005; Bashan et al. 2003).

What potential role would these early peptides have played? Some clues may surface from examination of ribosomal proteins in the structure of modern ribosomes, (Figure 1) where r-proteins are located at the interfaces of A-minor RNA-RNA interactions, namely Adenine and the minor groove of an RNA double helix (Nissen et al. 2001). It has been proposed that the bulk of the rRNA evolved *via* extensions of such A-minor interactions (Bokov et al. 2009) and it is more than likely that the initial proto-ribosome dimer was held together by similar RNA-RNA interactions. However, these interactions are mediated by only 4–6 hydrogen bonds, hence, for efficient function under changing environmental conditions over a significantly long period additional support can be advantageous. In the example of L2 and L36 in the 50S subunit of *D. Radiodurans*, the bulk of the RNA-protein interactions are mediated by positively charged residues. Interestingly, as with most ribosomal proteins, L2 and L36 are lysine, arginine and histidine rich, likely due to their greater propensity to form favorable complimentarity between their positive charge and the negatively charged phosphate ester backbone.

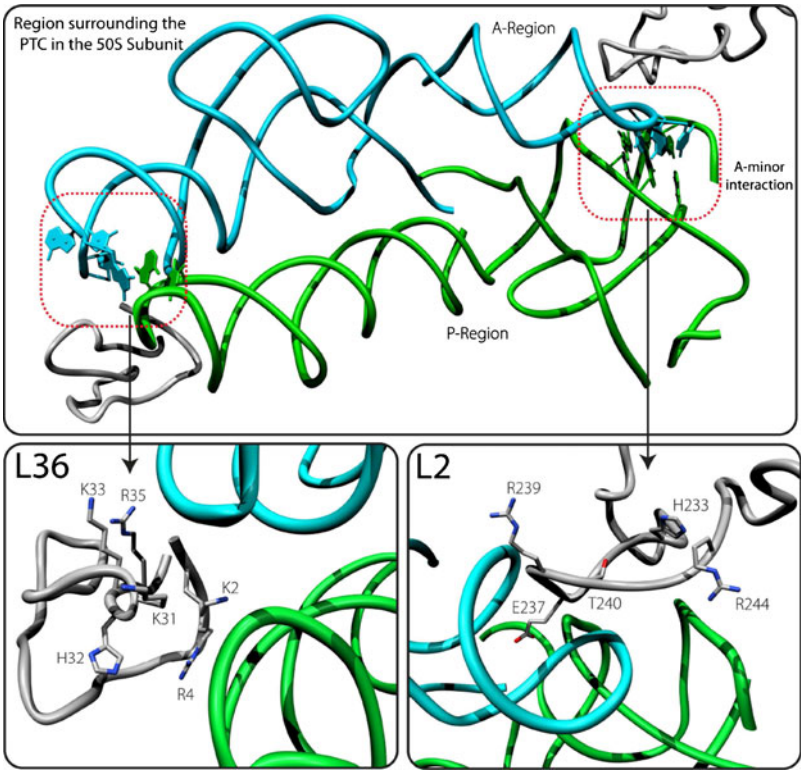
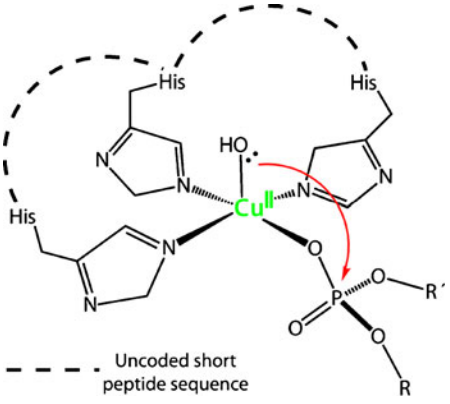


Fig. 1 A portion of the rRNA and proteins surrounding the peptidyl transferase centre in the 50S ribosomal subunit of *Deinococcus Radiodurans*, highlighting the stabilization of the A-minor interactions (in red box) by rProteins L2 and L36. Shown in the lower boxes is the RNA stabilized by positively charged amino acids side chains. The blue and green ribbons indicate the RNA backbone, the grey ribbons indicate the C-alpha positions of the proteins, with selected sidechains visible (PDB entry: 2ZJR) (Harms et al. 2008)

These clues on stabilizing role that ribo-proteins play, may indicate the driving force for the production of small peptides in the RNA world in accord with the observation that even a small arginine rich peptide portion of the Tat protein (a HIV-1 regulatory protein), bind with high selectivity and affinity with its cognate RNA partner, the transcription activation response element (TAR) (Sannes-Lowery et al. 1997).

In a world where replication, metabolism and chemistry is governed by RNA, one of the survival traits that emerging peptides must have exploited was binding to the agents of their production. The amino acids most likely to fulfill this role of RNA binding are ones that are positively charged (*e.g.* histidine, lysine and arginine), those heavily involved in modern RNA binding motifs. Furthermore, it has been postulated that it is these amino-acids that have the a 'higher catalytic propensity' (Kun et al. 2007), making them the ideal candidates for a dual role entities, either as naked peptides or as RNA-peptide hybrids. These types of amino-acids bear chemical functionality that makes them rich in reactivity. For example the imidazole ring in histidine can act as proton transfer agent, as well as a nucleophilic catalyst (Roth et al. 1998). Similarly, the guanidine in arginine has the ability to stabilize transition states of certain reactions such as phosphate ester cleavage (Kim et al. 1991). Conceivably, even short unfolded peptides rich in 'lysine', 'histidine' and 'arginine' could have served as simple catalysts or co-factors (Kun et al. 2007), yielding a selection advantage over peptides without these amino-acids. In support of this claim it has been shown that 'random' oligomers and polymers containing imidazole have catalytic properties for a range of chemical reactions (Okhapkin et al. 2004; Wulff 2002).

A mechanism that could have 'super-charged' the catalytic propensity of these proto-peptides would have been the formation of transition metal complexes involved in a larger range of chemical functionality and catalysis. For example short peptides, such as minimized constructs from avian prions and Amyloid-β peptides bind Cu(II) and Zn(II) ions with high affinity (Hornshaw et al. 1995; Syme et al. 2004), and Cu(II) and Zn(II) complexes that are coordinated by 'histidine' like moieties are fully functional catalysts for simple reactions such as manipulation of phosphate esters (Belousoff et al. 2008; Young et al. 1995) what would have been beneficial in an RNA world. Out of the pool of histidine rich peptides, it is possible that a certain amount could adopt a geometry enabling accommodation of transition metal ions, thus forming an 'active' site. A proposed example of a metallo-proto-peptide catalyst is shown in scheme 1, where a Cu(II) bound histidine rich peptide conjugate would be able to catalyze the hydrolysis of phosphate esters.



Scheme 1 Proto-peptide-metal complex as an early catalyst. A Copper(II)-histidine complex that could act as a potential phosphatase

We therefore propose that during the stage of uncoded translation preference for elongation may have been given to amino acids that were chemically compatible with the ribozyme reacting with them, as well as with other ribozymes and components in its immediate surroundings. Thus, suggesting preferred elongation of combinations of histidines, lysines and arginines in preference to other amino acids. A possible mechanism for such preference of the positively charged residues by the proto-ribosome could be the result of both substrate binding and product release. Positively charged amino acids were likely to fit well within the electrostatic field of the negatively charged RNA-made binding pocket. On the same principle, short peptides with large portion of positively charged amino acids could have spent more time in its surroundings before termination occurred. As at initial stages of the proto-ribosome evolution termination was likely to be a stochastic process, positively charged peptides probably had higher probability to be longer than neutral or acidic chains. Interestingly, phylogenetic analysis indicated that many of the last universal common ancestor (LUCA) peptide motifs are embedded within modern proteins involved in either RNA processing or nucleotide binding (Sobolevsky et al. 2007; Trifonov 2009; Trifonov et al. 2009). Moreover, LUCA motifs consistently contain a GKT peptide combination, supporting potential preference for that these amino acids (*i.e.* H, K, R) were the first to be incorporated into an early genetic code for coded protein production (Kun et al. 2007).

While the exact nature of the crossover from uncoded peptide elongation to the emergence of an amino-acid coding system is shrouded in mystery, evidence about how proteins interact with RNA in modern life can yield clues as to what sort of peptides were selected in an uncoded RNA world. If certain peptides were chemically selected and elongated, the pool of biological molecules available for the ribozymes to interact with becomes enriched with a smaller subset of amino acids or small peptides, concurrently opening the door for primitive coding mechanisms (such as coding coenzyme handles, Szathmary 1993) to emerge in the RNA world as well as providing rudimentary catalysts.

Acknowledgments
We thank all members of the ribosome group at the Weizmann Institute for continuous interest for fruitful discussions. Support was provided by the US National Inst. of Health (GM34360), and the Kimmelman Center for Macromolecular Assemblies. CD is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities. AY holds the Martin and Helen Kimmel Professorial Chair.

References

Agmon I, Bashan A, Zarivach R, Yonath A (2005) Symmetry at the active site of the ribosome: structural and functional implications. Biol Chem 386(9):833–844

Agmon I, Bashan A, Yonath A (2006) On ribosome conservation and evolution. Israel J Ecol Evol 52:359–374

Bashan A, Agmon I, Zarivach R, Schluenzen F, Harms J, Berisio R, Bartels H, Franceschi F, Auerbach T, Hansen HAS, Kossoy E, Kessler M, Yonath A (2003) Structural basis of the ribosomal machinery for peptide bond formation, translocation, and nascent chain progression. Mol Cell 11:91–102

Belousoff MJ, Tjioe L, Graham B, Spiccia L (2008) Synthesis, X-ray crystal structures, and phosphate ester cleavage properties of bis(2-pyridylmethyl)amine

copper(II) complexes with guanidinium pendant groups. Inorg Chem 47 (19):8641–8651

Bokov K, Steinberg SV (2009) A hierarchial model for evolution of 23S ribosomal RNA. Nature 457:977–980

Gluehmann M, Zarivach R, Bashan A, Harms J, Schluenzen F, Bartels H, Agmon I, Rosenblum G, Pioletti M, Auerbach T, Avila H, Hansen HA, Franceschi F, Yonath A (2001) Ribosomal crystallography: from poorly diffracting micro-crystals to high resolution structures. Methods 25:292–302

Harms JM, Wilson DN, Schluenzen F, Connell SR, Stachelhaus T, Zaborowska Z, Spahn CM, Fucini P (2008) Translational regulation via L11: molecular switches on the ribosome turned on and off by thiostrepton and micrococcin. Mol Cell 30:26–38

Hornshaw MP, McDermott JR, Candy JM, Lakey JH (1995) Copper binding to the N-therminal tandem repeat region of mammalian and avian prion protein: structural studies using synthetic peptides. Biochem Biophys Res Comm 214 (3):993–999

Kim EE, Wyckoff HW (1991) Reaction mechanism of alkaline phosphatase based on crystal structures: two-metal ion catalysis. J Mol Biol 218(2):449–464

Kun A, Pongor S, Jordan F, Szathmary E (2007) Catalytic propensity of amino acids and the origins of the genetic code and proteins. In: Barbieri M (ed) The codes of life: the rules of macroevolution, 1st edn. Springer

Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. Proc Natl Acad Sci USA 98(9): 4899–4903

Okhapkin IM, Bronstein LM, Makvaeva EE, Matveeva VG, Sulman EM, Sulman MG, Khokhlov AR (2004) Thermosensitive imidazole-containing polymers as catalysts in hydrolytic decomposition of p-nitrophenyl acetate. Macromolecules 37:7879–7883

Roth A, Breaker RR (1998) An amino acid as a cofactor for a catalytic polynucleotide. Proc Natl Acad Sci USA 93:4537–4542

Sannes-Lowery KA, Hu P, Mack DP, Mei H-Y, Loo JA (1997) HIV-1 Tat peptide binding to TAR RNA by electrospray ionization mass spectrometry. Anal Chem 69 (24):5130–5135

Sobolevsky Y, Frenkel ZM, Trifonov EN (2007) Combinations of ancestral modules in proteins. J Mol Evol 65:640–650

Syme CD, Nadal RC, Rigby SEJ, Viles JH (2004) Copper binding to the Amyloid-β (Aβ) peptide associated with Alzheimer's Disease. J Biol Chem 279(18):18169–18177

Szathmary E (1993) Coding coenzyme handles: a hypothesis for the origin of the genetic code. Proc Natl Acad Sci USA 90:9916–9920

Trifonov EN (2009) The origin of the genetic code and of the earliest oligopeptides. Res Microbiol. doi:10.1016/j.resmic.2009.05.004

Trifonov EN, Frenkel ZM (2009) Evolution of protein modularity. Curr Opin Struct Biol 2009 19:335–340

Wulff G (2002) Enzyme-like catalysis by molecularly imprinted polymers. Chem Rev 102 (1):1–25

Young MJ, Wahnon D, Hynes RC, Chin J (1995) Reactivity of copper(II) hydroxides and copper(II) alkoxides for cleaving an activated phosphate diester. J Am Chem Soc 117:9441–9447

E-mail: ada.yonath@weizmann.ac.il

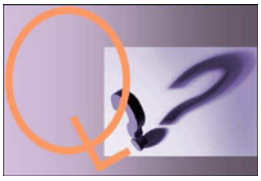
Workshop OQOL'09

Extended Abstracts for the Following Selected Question

- **Plausibility of the RNA World.**

Premise. The origin of life on the basis of a prebiotic family of RNAs is still a preferred scenario. This assumes, however, that RNA is formed prebiotically, while the question 'what made RNA?' is still unanswered. In fact, until now there is not even an accepted view of a robust prebiotic synthesis of mononucleotides, despite the considerable amount of work in the field by exquisite chemists. And, even if that would be discovered, still we would need to find a prebiotic way to couple the units in a 3′–5′ configuration to one another. And, finally, even if this also would be known, we would have to find out how a specific macromolecular sequence could be synthesized in many identical copies (see also the question above), to give a concentration of, say, 10^{-12}M in solution (which implies, in turn, more than 10^{13} (quasi) identical copies in one liter, or ca. 10^{7} identical copies in one microliter). One might conclude that the prebiotic synthesis of RNA is still a chimera from the scientific point of view.

The question. Do you share these arguments and rather bleak view? Which experiments or arguments can you suggest to counteract these objections against the "prebiotic" RNA world?



Excursions in an RNA World

Peter Schuster

Theoretical Biochemistry Group, University of Vienna, Austria

Keywords Allosteric Regulation • Exponential Growth • Neutral Networks • Origin of Chirality • Pyrimidine Ribonucleotide Synthesis • RNA Replication • RNA Switch • Suboptimal Structures

The question of the plausibility of an RNA world has two different aspects: (i) the historical dimension dealing with the problem how an RNA world could arise under prebiotic conditions, and (ii) the mechanistic dimension aiming at a comprehensive description of RNA properties and functions in a world without present day proteins. The crucial issue of the second aspect is whether or not an RNA only scenario is sufficiently rich in properties and functions in order to allow for further development into our present DNA-RNA-protein world. In the moment both approaches appear as 'work in progress', although our present knowledge on basic properties of RNA molecules seems to be closer to a principal understanding.

The historical aspect of the RNA world is not in the focus of this contribution but one comment is worth to be made: 'Never say never' in prebiotic chemistry. Two more or less arbitrarily chosen discoveries are mentioned that may initiate a modification of current research strategies: (i) the initial creation of homochirality and (ii) the prebiotic synthesis of pyrimidine nucleotides.

In 1953 the physical chemist Frederick Charles Frank (1953) published a possible mechanism for autocatalytic production of enantiopure products from achiral substances under non-equilibrium conditions. No experimentalist took Frank's model seriously until Kenso Soai reported an organic reaction that showed the predicted behavior (Soai et al. 1995). Although the Soai reaction is far away from any prebiotic processes, it complements the crystallization mechanism also leading to high enrichment of single enantiomers (Kondepudi et al. 1990). Chiral symmetry breaking is no longer mystery-like, still it is a long way to go from a single enantiopure compound to the chiral biological world.

The second new idea comes from nucleotide chemistry. The prebiotic production of pyrimidine ribonucleotides in the conventional way through joining phosphate, ribose and cytosine or uracil is notoriously difficult to visualize. John Sutherland and coworkers published a novel route to synthesize these compounds (Powner et al. 2009; Szostak 2009). Instead of linking the completed ring compounds they synthesize both rings around the already existing CN-bond between the ribose and the pyrimidine ring. Although it remains to be shown that this pathway is indeed plausible under prebiotic conditions, the novel reaction scheme introduces an alternative way of thinking about possible paths leading to the building blocks of ribonucleotides.

RNA is often considered a 'magic molecule' because every few years a new function is discovered that is based on its specific molecular properties (Gesteland et al. 2006). Despite the rich repertoire of catalytic properties of RNA molecules, until recently no assay was known that replicates longer stretches of RNA without the help of protein enzymes (Johnston et al 2003; Zaher and Unrau 2007). In particular, the goal to prepare a replicase ribozyme that can replicate itself in order to allow for the onset of evolution in the Darwinian sense is still in the distant future. Nevertheless, self-sustained replication of a cross-catalytic ensemble of two ribozymes (E,E′) has been recently reported (Lincoln and Joyce 2009). The two ribozymes act mutually as templates for production of the other ribozyme through ligation of two building blocks ($A+B \rightarrow E$; $A'+B' \rightarrow E'$). The E,E′ ensemble grows exponentially with an approximate doubling time of 1 h and the amplification can be continued to 'infinity' provided building blocks are supplied. The ribozymes contain variable regions where mutations have only minor effects and many functional building blocks can be constructed. A population of ribozymes was successfully designed to evolve by recombination, where E and E′ were assembled from a pool of A, A′, B and B′ modules. Further experiments showed coevolution of two ribozymes in a pool of five alternative substrates in the sense of exploiting two 'ecological' niches (Voytek and Joyce 2009).

Analysis of RNA (secondary) structure reveals a number of features that are important for an RNA world scenario (Schuster 2006): (i) neutrality in the sense that RNA sequence to function mappings are many to one, (ii) strong coupling of closely related selectively neutral sequences in replicating ensembles, and (iii) most RNA molecules have a rich spectrum of suboptimal conformations that play an important role in the regulation of multifunctional molecules. The search for function in RNA sequence space is largely facilitated by the fact that a large number of sequences share the same function and form neutral networks in sequence space (Schuster et al. 1994; Huang and Szostak 2003; Held et

al. 2003). The neutral selection scenario developed by Motoo Kimura (1983) is correct only for pairs of sequences with Hamming distances three or larger. Sequences of Hamming distance one appear in equal amounts in populations, sequences of Hamming distance two appear at a fixed ratio and, in other words, the close relatives coevolve and no selection takes place (Schuster and Swetina 1988). The mathematical result was confirmed recently by extensive computer simulations (Schuster P, 2009, unpublished results). Suboptimal and flexible conformations are important two different features of RNA function: (i) multiple (meta)stable states and (ii) conformational changes induced by allosteric binding of ligands. Provided the kinetic barriers for the interconversion of two or more structures are sufficiently high, molecules may exist in one stable and one or more metastable conformations. An example of a designed RNA switch with two different structures and its sensitivity against mutation was recently reported (Nagel et al 2006). Another designed example seems to be highly relevant for early RNA world scenarios: The structures of two different ribozymes of the same chain length but completely different functions and sequences were used to design an 'intersection sequence' that can fold into both structures. The synthesized molecule with the intersection sequence forms the two different structures and shows both catalytic activities (Schultes and Bartel 2000). Moreover it can be transformed by single point mutations and base pair exchanges along neutral paths (Held et al. 2003) through sequence space into the reference sequences without loosing catalytic activity.

RNA switches are important regulatory elements for metabolic processes (Mandal and Breaker 2004). Because of the structural simplicity and the relative easiness of their design (Tang and Breaker 1997) they are excellent candidates for prebiotic regulation of function. The principle of these allosterically controlled ribozymes—called aptazymes—is very simple: A flexible part of the RNA molecule rigidifies on binding of a ligand and this converts an inactive RNA molecule into an active ribozyme. Riboswitches control, for example, the synthesis of enzymes involved in the production of metabolites. When the metabolite is present it binds to the mRNA and translation stops, when the concentration of the metabolite falls below a certain threshold the ligand dissociates from the mRNA, which changes conformation and becomes translated, and the metabolite is produced again.. An allosteric switch has also been introduced into the self-replicating ensemble of two ribozymes (E,E′) mentioned above (Lam and Joyce 2009). Exponential amplification of RNA sets in when the ligand of the aptazyme is present and it stops in absence of the regulatory molecule.

References

Frank FC (1953) On spontaneous asymmetric synthesis. Biochim Biophys Acta 12:459–463

Gesteland RF, Cech TR, Atkins JF (eds) (2006) The RNA World. Third Ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Held DM, Greathouse ST, Agrawal A, Burke DH (2003) Evolutionary landscapes for the acquisition of new ligand recognition by RNA aptamers. J Mol Evol 57:299–308

Huang K, Szostak JW (2003) Evolution of aptamers with a new specificity and new secondary structure from an ATP aptamer. RNA 9:1456–1463

Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP (2003) RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension. Science 292:1319–1325

Kimura M (1983) The neutral theory of molecular evolution. Cambridge, UK: Cambridge University Press

Kondepudi DK, Kaufman RJ, Singh N (1990) Chiral symmetry-breaking in sodium-chlorate crystallization. Science 250:975–976
Lam BJ, Joyce GF (2009) Autocatalytic aptazymes enable ligand-dependent exponential amplification of RNA. Nature Biotech 27:288–292
Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. Science 323:1229–1232
Mandal M, Breaker RR (2004) Gene regulation by riboswitches. Nature Rev Mol Cell Biol 5:451–463
Nagel JHA, Flamm C, Hofacker IL, Franke K, de Smit MH, Pleij CWA (2006) Structural parameters affecting the kinetic competition of RNA hairpin formation. Nucleic Acids Res 34:3568–3576
Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyridine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242
Schultes EA, Bartel DB (2000) One sequence, two ribozymes: Implications for the emergence of new ribozyme folds. Science 289;448–452
Schuster P (2006) Prediction of RNA secondary structures: From theory to models and real molecules. Rep Prog Phys 69:1419–1477
Schuster P, Fontana W, Stadler PF, Hofacker IL (1994) From sequences to shapes and back: A case study in RNA secondary structures. Proc Roy Soc London B 255:279–284
Schuster P, Swetina J (1988) Stationary mutant distributions and evolutionary optimization. Bull Math Biol 50:635–660
Soai K, Shibata T, Morioka H, Choji K (1995) Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. Nature 378:767–768
Szostak JW (2009) Systems chemistry on early Earth. Nature 459:171–172
Voytek SB, Joyce GF (2009) Niche partitioning in the coevolution of two distinct RNA enzymes. Proc Natl Acad Sci USA 106:7780–7785
Tang J, Breaker RR (1997) Rational design of allosteric ribozymes. Chem&Biol 4:453–459
Zaher WK, Unrau PJ (2007) Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. RNA 13:1017–1026

E-mail: pks@tbi.univie.ac.at

On the Likelihood of the Abiotic Formation of RNA Molecules

Robert Shapiro

Department of Chemistry New York University, USA

Keywords RNA World • Total Synthesis • Prebiotic Simulation • Cyanoacetylene • Nucleotides

The publication by Walter Gilbert which coined the phrase “RNA world” called for the assembly of the first RNAs from “a nucleotide pool” (Gilbert 1986). The nucleotide components of RNA are substances of considerable chemical complexity, bearing four chiral centers, and the regiospecific connection of the furanose form of the sugar ribose to a particular place on each of four heterocycle entities. If the nucleotides of the pool are to be capable of extensive polymerization, each must bear its phosphate on the same hydroxyl group of the ribose ring.

No driving force of nature or set of chemical circumstances is known which would produce the four RNA nucleotides without producing a host of similar substances, with alternative sugars bound to an extensive group of nitrogenous substances, and phosphate (or some substitute) connected at a variety of positions (Shapiro 2000). This array of N-glycosides should be accompanied by an even larger concentration of simpler substances, such as alkyl phosphates, which would serve to terminate chains in any process that favored polymerization. The "progress" that has been made by skilled chemists in this area more reflects an achievement in the total laboratory synthesis of RNA rather than any recapitulation of events on the early Earth.

Much confusion has been caused in journals and the media by the mistaken assumption that these two processes are equivalent. As an example, the recent noteworthy total synthesis of cytidine 2′,3′-phosphate by Powner et al. (2009) can be cited. Their route was efficient and elegant in that it employed only four organic chemicals of modest size, and inorganic phosphate, in a limited number of steps that under optimized conditions afforded good yields of the desired product.

As conducted, this effort matched the definition put forth by Nicolaou and Sorensen (1996): "*Total synthesis* is the chemical synthesis of a molecule, usually a natural product, from relatively simple starting materials…The ultimate goal of organic synthesis is to assemble a given organic compound …in the most efficient way… The science of organic synthesis is constantly enriched by new inventions and discoveries pursued deliberately for their own sake or as subgoals within a program directed toward the synthesis of a target molecule."

This practice should be distinguished from that of prebiotic simulation, in which a mixture of chemicals thought to be present on the early Earth is allowed to interact without a predetermined agenda, for example the classic Miller-Urey spark discharge experiment (Miller, 1953). Unfortunately, the Powner, et al. paper made the claim "The starting materials of the synthesis—cyanamide, cyanoacetylene, glycolaldehyde, glyceraldehyde and inorganic phosphate—are plausible prebiotic feedstock molecules, and the conditions of the synthesis are consistent with potential early-Earth geochemical models."

A full analysis of this claim would go beyond the limits of this article. However it can be noted that their first reaction contained 1 M concentrations of only three substances: cyanamide, glycolaldehyde and phosphate. These concentrations are orders of magnitude greater than those usually estimated for chemicals in the early oceans (for references, see Shapiro, 1999). Many other substances whose presence in many environments on the early Earth could be expected; formaldehyde, cyanide, amines and amino acids, for example, were excluded from the reaction mix. A later reaction in the synthesis employed cyanoacetylene in a concentration of 0.49 molar. The plausibility of this substance as a "prebiotic feedback molecule" was justified only by the citation of two papers. One tabulated its astronomical detection (with over 130 other molecules) in the interstellar gas (Thaddeus 2006) and the other reported its transient formation by the action of an electric discharge in a methane-nitrogen atmosphere, and subsequent reaction with a variety of chemicals (Sanchez et al. 1966). A more extensive discussion of the reaction of cyanoacetylene with a variety of simple nucleophiles has been published with this author (Shapiro 1999). In the absence of data from an extensive series of competition experiments, it seems unwise to assume that cyanoacetylene would have accumulated to any extent on the early Earth. The reaction sequence of Powner, et al., while admirable as another example of well-executed total synthesis, must yet be judged as an extremely unlikely representation of geochemical events before life began.

Difficulties of this type, and a number of other problems, caused Gerald Joyce and Leslie Orgel (2006) to declare that the abiotic formation of RNA would constitute a "near miracle". (Many other chemists, myself included, agree with this assessment). If we reject the idea that RNA, and other information-rich biopolymers were present at the start of life, then we arrive at an alternative, satisfactory solution for their origin. RNA first appeared through natural selection in living organisms, as the result of an extensive series of events, each of which had its own justification.

If we accept this argument, then we must conclude that the earliest forms of life functioned through the activities of sets of smaller, abiotically available molecules. To understand the origin of life, we must understand and if possible model these processes.

References

Gilbert W (1986) Origins of life: the RNA world. Nature 319:618

Joyce GF, Orgel LE (2006) Prospects for understanding the origin of the RNA world. In: Gesteland RF, Cech TR, Atkins JF (eds) The RNA World 3rd edn. Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY, pp 23–56

Miller SL (1953) A production of amino acids under possible primitive earth conditions. Science 117:528–529

Nicolaou KC, Sorenson EJ (1996) Classics in Total Synthesis. VCH Publishers, New York

Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242

Sanchez RA, Ferris JP, Orgel LE (1966) Cyanoacetylene in prebiotic synthesis. Science 154:784–785

Shapiro R (2000) Prebiotic cytosine synthesis: a critical analysis and implications for the origin of life. Proc Natl Acad Sci USA 96:4396–4401

Shapiro R (2000) A replicator was not involved in the origin of life. IUBMB Life 49:173–176

Thaddeus P (2006) The prebiotic molecules observed in the interstellar gas. Phil Trans R Soc B 361:1681–1687

E-mail: rs2@nyu.edu

Linking the RNA World to Modern Life: The Proto-Ribosome Conception

Anat Bashan, Matthew J. Belousoff, Chen Davidovich, Ada Yonath

Dept of Structural Biology, Weizmann Institute, Israel

Keywords Proto-ribosome • Peptide Bond Formation • Genetic-code Evolution

In order to tackle the plausibility of the 'RNA world' hypothesis we aimed at scrutinizing an ancient self-folded RNA entity that functioned in the RNA world and evolved into a simple machine capable of catalyzing chemical reactions, including the formation of peptide bonds. By exploiting structural, biochemical, computational and modeling experiments, the remnant of this machine was identified within the contemporary ribosome (Baram and Yonath 2005; Agmon et al. 2006; Agmon et al. 2009; Davidovich et al. 2009 Yonath 2009).

The ribosome is the universal multi-component macromolecular cellular assembly that decodes the genetic information and efficiently elongates nascent polypeptide chains under the mild conditions of the modern life. The contemporary ribosomes are ribonucleoprotein assemblies, with molecular weights of 2.5 and 4MDa (for prokaryotes and eukaryotes, respectively). Despite this significant differences in their sizes, the core functional regions of ribosomes from all kingdoms of life exhibit remarkable conservation. Among them is a universal symmetrical 'pocket-like' substructure, an extraordinary feature in the otherwise asymmetric ribosome. It is composed of 180 ribosomal RNA (rRNA) nucleotides (Bashan et al. 2003 MC; Agmon et al. 2005), and hosts the peptidyl transferase center (PTC) namely the site of peptide bond formation. This symmetrical region provides the framework for the positioning of the ribosomal substrates in favorable stereochemistry for peptide bond formation and for substrate-mediated catalysis (Bashan 2003 et al. MC; Gregory et al. 2004; Bieling et al. 2006; Weinger et al. 2007; Bashan and Yonath 2008). Furthermore, by encircling the PTC the architecture of this region confines the void required for the motions involved in tRNA 3′ end translocation, required for the successive peptide bond formations, thus enabling the ribosome polymerase activity (Bashan et al. 2003; Agmon et al. 2005).

The ribosomal symmetrical region seems to be preserved throughout evolution (Agmon et al. 2006; Agmon et al. 2009; Davidovich et al. 2009; Yonath 2009) and the fold of each of its halves resembles the main building block of "ancient" as well as "modern" functional RNA molecules of comparable size (e.g. gene regulators, riboswitches, ribozymes catalyzing the phosphodiester cleavage, RNA processors etc), hence suggesting that it could have existed in the RNA world as a self folded autonomous entity. This entity could have functioned as an apparatus catalyzing various reactions involved in RNA metabolism, as well as peptide bond formation and non-coded oligopeptides elongation.

Support for the existence of RNA entity capable of self replication, folding and dimerization are the recent non-enzymatic synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions (Powner et al. 2009; Szostak 2009) and the demonstration that RNA oligomers can be obtained non-enzymatically from activated RNA precursors (Pino et al. 2008; Krzyaniak et al. 1994). The dimerization in a symmetrical manner of self folded motifs of identical, similar or different sequences, may have occurred spontaneously, resulting occasionally in 'pocket-like' structures capable of hosting spontaneously produced amino-acids conjugated with single or short oligonucleotides, (Ilangasekare et al. 1995; Giel-Pietraszuk et al. 2006; Lehmann et al. 2008) serving as substrates for peptide bond formation.

The more stable constructs of these 'pocket-like' molecular dimers might have survived under various environmental conditions. Among them, those that would accommodate suitable substrates at the appropriate stereochemistry enabling peptide bond formation have been evolutionarily favored. As it is assumed that in the prebiotic era RNA chains could self replicate (Eigen 1993; Smith and Szathmáry 1995; Lincoln and Joyce 2009; Woese 2001; Yarus 2002), it is conceivable that phenotypes with favorable properties could have been synthesized in many copies. It is likely that some of these phenotypes were originated by fusion of two different or duplication of two identical sequences, resembling gene elongation events (Fani and Fondi 2009).

Based on the high conservation of the ribosomal region assigned as the proto-ribosome and on its capability to provide all of the architectural elements required for

amino acid oligomerization, we assumed that it existed in the RNA world and functioned in a fashion similar its precedent within to the contemporary ribosome. The hypothesis of a self assembled ribosomal active site, which is still implanted in the internal core of the modern ribosome, triggered biochemical experiments aimed at revealing the tendency for self folding and dimerization of RNA chains. These yielded biochemical evidence supporting the existence of a dimeric proto-ribosome, and provided hints for a feasible pathway for acquiring the structural elements necessary for coded amino acid polymerization. Hence shedding light on the emergence of the contemporary genetic translation apparatus from rather short RNA oligomers (Fig. 1).

We found that some, albeit not all, RNA chains with sequences resembling those observed in the current ribosome, are capable of forming dimers that may adopt a 'pocket-like' structure (Davidovich et al. 2009). Furthermore by site-directed mutagenesis we showed that the tendency for dimerization, a prerequisite for obtaining the catalytic centre, is linked to the fold of the proto-ribosome two components, thus indicating that functional selection at the molecular level existed already in the prebiotic era. Consistently, it is conceivable that 'pocket-like' RNA entities were assembled spontaneously from a pool of RNA chains.

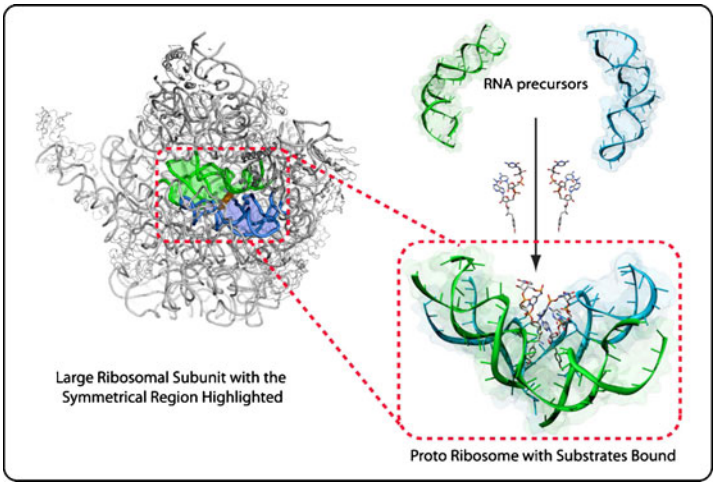


Fig. 1 Left: the symmetrical region at the heart of the large ribosomal unit. Its two halves are colored in blue and green and actual peptide bond formation site is shown in red.
Right: Top: the precursors of the proto-ribosomes in their assumed conformation Bottom: the pocket-like entity resulting from the dimerization of the two precursors

Among the products of these early amino acid elongation processes, those molecular entities possessing central, albeit primitive catalytic and/or synthetic properties, became the templates for enhanced production (see Belousoff et al. this issue), survived evolution pressures, and underwent natural selection. Among the key tasks performed by the initial oligopeptides is stabilizing the proto-ribosome and/or other components confined in its surrounding, within assemblies that could evolve into "proto-cells". As it is likely that

subsequently the proto-ribosomes underwent optimization from non-genetic peptide bond formation towards performing genetically driven translation, it is conceivable that the ancient proto-ribosome in its functionally-optimized version is still embedded in the core of the modern ribosome, and that the symmetrical region of the modern ribosome originated from the proto-ribosome.

In short, here we present structural tools for investigating possible pathways in the evolution of modern life and approaching key questions, such as: Did the ancient translation apparatus survive selection pressure? Does its relic reside within the modern ribosome? What was the evolution conduit leading to its successive optimization?

Acknowledgments
We thank Ilana Agmon and all members of the ribosome group at the Weizmann Institute for continuous interest for fruitful discussions; Ofir Sade-Falk and Leena Taha for excellent technical assistance. Support was provided by the US National Inst. of Health (GM34360), and the Kimmelman Center for Macromolecular Assemblies. CD is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities. AY holds the Martin and Helen Kimmel Professorial Chair.

References

Agmon I, Bashan A, Zarivach R, Yonath A (2005) Symmetry at the active site of the ribosome: structural and functional implications. Biol Chem 386:833–844

Agmon I, Bashan A, Yonath A (2006) On ribosome conservation and evolution. Isr J Ecol Evol 52:359–379

Agmon I, Davidovich C, Bashan A, Yonath A (2009) Identification of the prebiotic translation apparatus within the contemporary ribosome. http://precedingsnaturecom/documents/2921/version/1

Baram D, Yonath A (2005) From peptide-bond formation to cotranslational folding: dynamic, regulatory and evolutionary aspects. FEBS Lett 579:948–954

Bashan A, Yonath A (2008) Correlating ribosome function with high-resolution structures. Trends Microbiol 16:326–335

Bashan A, Agmon I, Zarivach R, Schluenzen F, Harms J, Berisio R, Bartels H, Franceschi F, Auerbach T, Hansen HAS et al (2003) Structural basis of the ribosomal machinery for peptide bond formation, translocation, and nascent chain progression. Mol Cell 11:91–102

Bieling P, Beringer M, Adio S, Rodnina MV (2006) Peptide bond formation does not involve acid-base catalysis by ribosomal residues. Nat Struct Mol Biol 13:424–428

Davidovich C, Belousoff M, Bashan A, Yonath A (2009) The evolving ribosome: from non-coded peptide bond formation to sophisticated translation machinery. Res Microbiol 18:18

Eigen M (1993) The origin of genetic information: viruses as models. Gene 135:37–47

Fani R, Fondia M (2009) Origin and evolution of metabolic pathways. Phys Life Rev 6:23–52

Giel-Pietraszuk M, Barciszewski J (2006) Charging of tRNA with non-natural amino acids at high pressure. Febs J 273:3014–3023

Gregory ST, Dahlberg AE (2004) Peptide bond formation is all about proximity. Nat Struct Mol Biol 11:586–587

Illangasekare M, Sanchez G, Nickles T, Yarus M (1995) Aminoacyl-RNA synthesis catalyzed by an RNA. Science 267:643–647
Krzyzaniak A, Barciszewski J, Salanski P, Jurczak J (1994) The non-enzymatic specific amino-acylation of transfer RNA at high pressure. Int J Biol Macromol 16:153–158
Lehmann J, Reichel A, Buguin A, Libchaber A (2007) Efficiency of a self-aminoacylating ribozyme: effect of the length and base-composition of its 3′ extension. RNA 13:1191–1197
Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. Science 323:1229–1232
Maynard-Smith J, Szathmary E (1997) The major transitions in evolution. Oxford Uni Press, NY
Pino S, Ciciriello F, Costanzo G, Di Mauro E (2008) Nonenzymatic RNA ligation in water. J Biol Chem 283:36494–36503
Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242
Szostak JW (2009) Origins of life: systems chemistry on early Earth. Nature 459:171–172
Weinger JS, Strobel SA (2007) Exploring the mechanism of protein synthesis with modified substrates and novel intermediate mimics. Blood Cells Mol Dis 38:110–116
Woese CR (2001) Translation: in retrospect and prospect. RNA 7:1055–1067
Yarus M (2002) Primordial genetics: phenotype of the ribocyte. Annu Rev Genet 36:125–151
Yonath A (2009) Large facilities and the evolving ribosome, the cellular machine for genetic-code translation. J R Soc Interface online

E-mail: ada.yonath@weizmann.ac.il

A Model for the Modular Evolution of RNA Addressing Open Questions on the Origin of Life

Carlos Briones, Michael Stich, Susanna C. Manrubia

Centro de Astrobiología (CSIC-INTA), Madrid, Spain

Keywords RNA World • Modular Evolution • Functional Complexity • RNA Ligation • Hairpin Ribozyme • RNA Replicase

A main unsolved problem in the RNA world scenario for the origin of life is how a template-dependent RNA polymerase ribozyme emerged from short RNA oligomers generated by random polymerization of ribonucleotides (Joyce and Orgel 2006). Current estimates establish a minimum size about 165 nt long for such a ribozyme (Johnston et al. 2001), a length three to four times that of the longest RNA oligomers obtained by random polymerization on clay mineral surfaces (Huang and Ferris 2003, 2006). To overcome this gap, we have developed a stepwise model of ligation-based, modular evolution of RNA (Briones et al. 2009) whose main conceptual steps are summarized in Figure 1. This scenario has two main advantages with respect to previous hypotheses put forward for the origin of the RNA world: i) short RNA

modules resulting from template-independent polymerization on different microenvironments might suffice to produce the first functional RNAs in the absence of template replication; ii) modular evolution shortens adaptation times and generates complex structures that could not be directly selected. Therefore, ligation-based modular evolution might have bridged the gap between the last stages of the pre-RNA world and a fully established RNA world. The emergent information-based molecular machinery could subsequently evolve and be inherited by DNA-based precellular systems leading to the progenote. Although mainly focused on the origin of the RNA world, our model addresses in different ways several questions posed in the OQOL'09 workshop, as we discuss in the following.

Contingency vs. determinism—Our stepwise model describes levels of increasing biochemical complexity. Experimental and computational data support that the appearance of an RNA world may be the plausible outcome of molecular evolution even in the absence of template replication, and hence not a completely contingent event. The existence of many less RNA structural families than possible sequences points to a principle of 'canalized contingency'. This concept is analogous to the existence of attraction basins of sequences, where the initial randomness is strongly suppressed by the convergence to a limited set of structures. Determinism is related to the partial independence of the functional level from the microscopic level. We have shown in silico that the fraction of short, random molecules displaying catalytic activity is large enough to trigger the processes that lead to a modular origin of the RNA world.

Plausibility of the RNA world—One main problem stated in this question is the difficulty of obtaining many identical copies of a specific macromolecular sequence, apparently a requirement for the appearance of effective chemical function. However, different RNA sequences fold into identical structures, which group themselves into a reduced number of structural families. Since simple biochemical functions can be performed by slightly different structures of a given family, a pool of random, short oligomers may be a viable starting point for the RNA world. Assuming that there is a unique sequence able to perform a given function is misleading, since this premise overlooks the huge—and convenient—degeneration of the sequence-structure-function map in RNA.

Life as a unity or confederacy—Our scenario illustrates a possibly general principle for attaining functional complexity from short and non-informative molecular modules, ready to be combined in a constructive way. It suggests that the emergence of complex replicative molecules—the substrate of life—is due to a confederacy of subsystems which might have previously undergone partly independent evolution and selection. Once ligated into multi-modular molecules, the initial setting, characterized by the competition among modules, turned into a cooperative framework where the joint modules constituted selectable units. Analogous changes from confederacy to unity could have been further produced, for example by the eventual encapsulation of two polymerase ribozymes into a common compartment. From those transitions onwards, the evolution of the joined systems could erase the features of their prior, independent evolution, thus raising the question of whether, even if life started as a confederacy of modular units, this origin could be traced in modern cells.

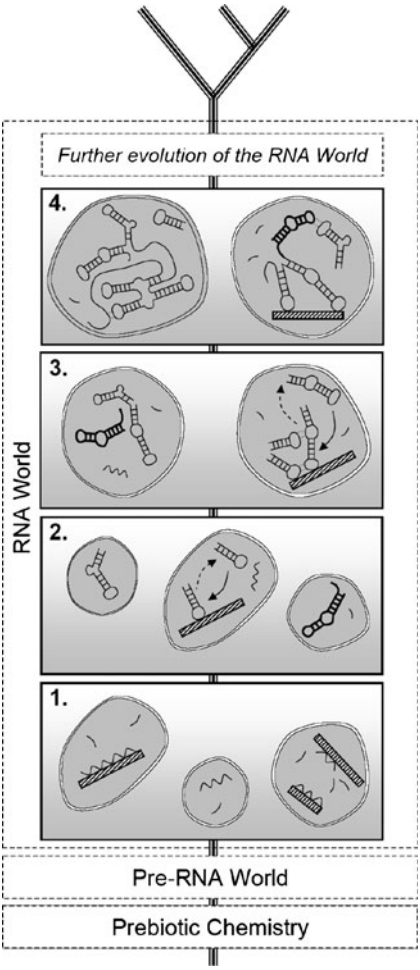


Fig. 1. Stepwise model for the modular evolution of RNA in the origin of life. The proposed evolutionary process can be divided into four conceptual steps, triggered once prebiotic chemistry had provided the required monomers and oligomerization was possible in the pre-RNA world. **1.** The abiotic polymerization of RNA from activated nucleotides could have occurred on mineral surfaces exposed to bulk solution (Huang and Ferris 2003, 2006) or within vesicles (Hanczyc et al. 2003), yielding up to 30–50mer random RNA oligomers. **2.** Every sequence folded into its minimum free energy structure. Computational analyses of the structural repertoire present in large populations (10^8 molecules) of random RNA sequences of length 12 to 40 nt reveal that topologically simple modules are the most abundant ones, especially hairpin structures and stem-loops (Stich et al. 2008). **3.** A fraction of hairpin modules could have displayed RNA ligase activity (in bold line), as certain ribozymes currently do, and thus catalyzed the assembly of larger, eventually functional molecules. Ligation processes allow a fraction of the combined molecules to retain their previous modular structure, such that structural and functional complexity can progressively increase even in the absence of template replication (Manrubia and Briones 2007). **4.** The iteration of that process could have assembled RNA molecules endowed with novel functionalities, paving the way to the emergence of a—relatively long and complex—ribozyme (Johnston et al. 2001) with

template-dependent polymerase activity. At this step, information-driven evolution would be triggered. More details on the model can be found in Briones et al. (2009).

Defining the very origin of life—A question indirectly addressed in the previous paragraph is the possibility that the first complex molecule required to establish a robust RNA world (i.e., an RNA polymerase) could have been the result of the interaction among a collection of modular subsystems, each with its own dynamics and proto-metabolic network of interactions. If simple modules are seen as functional and relatively stable entities, the scenario devised could also act as a meeting point for two traditionally opposite views: metabolism-first and information-first scenarios. This is, indeed, one of the most productive controversies behind the definition of the origin of life.

Is life an emergent property?—Once a functional ligase ribozyme opened the possibility of combining modules with different functionalities, new molecules endowed with unexpected chemical properties and activities could emerge. The resulting products are larger molecules whose functionality cannot be known a priori given the properties of the constituting modules. This fact has been experimentally demonstrated by in vitro evolution of RNA, where ligation and exchange of structural domains can be used to engineer new functional RNAs (reviewed in Joyce 2004). In our model, the appearance of a ligase molecule in a random pool of oligomers, the acquisition of new functionalities through modular evolution, and the origin of an RNA replicase, all represent emergent properties that could, in turn, be seen as pre-requisites for the emergence of life.

References

Briones C, Stich M, Manrubia SC (2009) The dawn of the RNA world: Towards functional complexity through ligation of random RNA oligomers. RNA 15:743–749

Hanczyc MM, Fujikawa SM, Szostak JW (2003) Experimental models of primitive cellular compartments: encapsulation, growth, and division. Science 302:618–622

Huang W, Ferris JP (2003) Synthesis of 35–40 mers of RNA oligomers from unblocked monomers. A simple approach to the RNA world. Chem Commun 12:1458–1459

Huang W, Ferris JP (2006) One-step, regioselective synthesis of up to 50-mers of RNA oligomers by montmorillonite catalysis. J Am Chem Soc 128:8914–8919

Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP (2001) RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science 292:1319–1325

Joyce GF (2004) Directed evolution of nucleic acid enzymes. Annu Rev Biochem 73:791–836

Joyce GF, Orgel LE (2006) Progress towards understanding the origin of the RNA world. In: Gesteland RF, Cech TR, Atkins JF (eds) The RNA world, 3rd edn. Cold Spring Harbour Lab., New York, pp 23–56

Manrubia SC, Briones C (2007) Modular evolution and increase of functional complexity in replicating RNA molecules. RNA 13:97–107

Stich M, Briones C, Manrubia SC (2008) On the structural repertoire of pools of short, random RNA sequences. J Theor Biol 252:750–763

E-mail: brioneslc@inta.es

The Prebiotic Synthesis of RNA and Pre-RNA

Jim Cleaves

Carnegie Institution for Science, Washington D.C, USA

Keywords Pre-RNA prebiotic synthesis RNA TNA • GNA

The idea has been put forth that life originated from self-replicating, catalytic RNA molecules that were formed spontaneously on the surface of the Earth via prebiotic chemistry (Orgel 1987). The idea is particularly attractive due to the preponderance of catalytic RNA in modern cells and the central role of RNA in translation (Cech et al 1981; Benner et al 1989; Gesteland et al 1999). Indeed this would be in many respects the simplest solution to the question of the origin of life on Earth.

However, after almost 50 years of laboratory effort, the prebiotic synthesis of RNA has proven rather difficult (Fuller et al 1972; Shapiro 1988), although there have been recent elegant model syntheses (Powner et al 2009). One solution to this possible dilemma is that RNA was preceded by a simpler genetic molecule whose prebiotic synthesis was considerably more facile (Joyce et al 1987). This is probably one of the most experimentally testable ideas in the study of the origins of life.

It is now known that there are a wide variety of molecules capable of Watson-Crick type base-pairing based on other sugars, other linkers, and other bases (Egholm et al 1992; Eschenmoser 2004; Zhang et al 2005; Bean et al 2006). Relatively few of these have received the focused attention RNA has with respect to abiological synthesis (Nelson et al 2000; Cleaves 2002), partly for the reasons cited above.

We have systematically studied the chemical properties of a number of nucleic acid analogues over the years, with an emphasis on robustness of prebiotic synthesis and stability. Among the analogues investigated are those shown in Fig. 1.

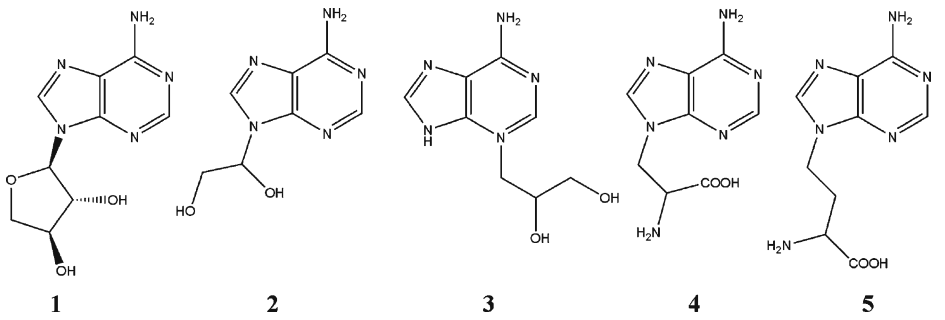


Fig. 1 Structures investigated

We have found that monomers **1** and **2** are likely as feasible prebiotically as ribonucleosides with respect to synthesis, and are of comparable stability to ribonucleosides over a wide range of pH values and temperatures, while monomers **3**, **4** and **5** decomposed remarkably quickly at even moderate temperatures (~40–60°C).

Such criteria should only be considered in the context of other important factors such as base-pairing ability and polymer stability. Nevertheless, it is apparent that there remain

many nucleic acid analogue structures to be investigated and compared with respect to modern biological nucleic acids.

References

Bean HD, Anet FA, Gould IR, Hud NV (2006) Glyoxylate as a backbone linkage for a prebiotic ancestor of RNA. Orig Life Evol Biosph 36:39–63

Benner SA, Ellington AD, Tauer A (1989) Modern metabolism as a palimpsest of the RNA world. Proc Nat Sci USA 86:7054–7058

Cech TR, Zaug AJ, Grabowski PJ (1981) In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. Cell 27:487–496

Cleaves HJ (2002) The reactions of nitrogen heterocycles with acrolein: scope and prebiotic significance. Astrobiol 2:403–415

Egholm M, Buchardt O, Nielsen PE, Berg RH (1982) Peptide nucleic acids (PNA) oligonucleotide analogues with an achiral peptide backbone. J Am Chem Soc 114:1895–1897

Eschenmoser A (2004) The TNA-family of nucleic acid systems: properties and prospects. Orig Life Evol Biosph 34:277–306

Fuller WD, Sanchez RA, Orgel LE (1972) Solid-state synthesis of purine nucleosides. J Mol Evol 1:249–257

Gesteland R, Cech T, Atkins J (1999) The RNA world, 2 Edn [Cold Spring Harbor Monogr. Ser. 37]

Joyce G, Schwartz A, Miller S, Orgel L (1987) The case for an ancestral genetic system involving simple analogs of the nucleotides. Proc Nat Acad Sci USA 84:4398–4402

Nelson K, Levy M, Miller S (2000) Peptide nucleic acids rather than RNA may have been the first genetic molecule. Proc Nat Acad Sci USA 97:3868–3871

Orgel LE (1987) The origin of self-replicating molecules, in self-organizing systems: the emergence of order. Plenum, New York

Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242

Shapiro R (1988) Prebiotic ribose synthesis: a critical analysis. Orig Life Evol Biosph 18:71–85

Zhang L, Peritz A, Meggers E (2005) A simple glycol nucleic acid. J Am Chem Soc 127:4174–4175

E-mail: hjcleaves@ciw.edu

Commentary on "Plausibility of an RNA World"

Addy Pross

Department of Chemistry, Ben-Gurion University of the Negev, Israel

Keywords RNA-World • Metabolic Cycles • Dynamic Kinetic Stability • Origin of Life

The question whether life originated through the emergence of some autocatalytic metabolic reaction cycle—the metabolism view (Kauffman 2000; Shapiro 2006; Segre et al. 2000; Morowitz et al. 2000; Wachtershauser 1988), or a self-replicating oligomer of variable sequence—the RNA-World view (Gesteland et al. 1999; Joyce 2002; Orgel 1998), remains controversial and a source of continuing debate, though it should be noted that

more recent considerations have indicated that a sharp demarcation between the two approaches may not be warranted (Eschenmoser 2007).

In a recent paper (Pross 2009) we have argued that by building on the concept of dynamic kinetic stability (Pross 2004, 2005), the two stages in life's establishment on earth—emergence and evolution, may be conceptually unified and considered as one continuous process governed by a single driving force principle—the drive toward greater dynamic kinetic stability. That view is strengthened by the realization that Darwinian theory has its roots in chemical kinetic theory, lending weight to the idea that biological phenomena on the one hand, and replicative chemical phenomena on the other, share a common underlying physicochemical basis. Indeed, the observation of Darwinian natural selection at the molecular level (Mills et al. 1967, Eigen 1992) is a striking illustration of the intimate chemistry-biology interrelation. Accordingly, we believe insights into the relatively uncertain and poorly understood prebiotic emergence phase can be obtained by relating it to the later, relatively well understood evolutionary phase. Simply put, ***evolutionary patterns gleaned from evolutionary biology may provide insights into the chemistry of emergence***. Let us now consider possible applications of this way of reasoning to further probe the nature of the primal replicator—metabolic or genomic?

Modern Darwinian thinking endorses the following two central ideas. First, all living systems, whether cyanobacteria, believed to have existed on earth for some 3.5 billion years, or more recent life forms such as we humans, utilize the same basic nucleic acid genomic system. Second, evolution is considered to have taken place by small incremental steps, which in molecular terms is attributed to genome sequence mutation. If we now build on our proposal of utilizing the pattern observed in evolution as a likely model for describing the process of emergence, we are led to several conclusions. First, both the ***universality*** and the ***stability*** of the genomic system of information storage over the billions of years during which life on earth evolved, suggest that the mode of information storage in the prebiotic phase would have been similar in kind, namely, one based on an oligomeric genomic system, rather than one based on a non-genomic metabolic system. That way of thinking in itself lends support to the RNA-world view, in which the emergence of some genomic RNA-like replicator was the primordial event that led to the emergence of life. However this argument in support of a genomic origin may be taken a step further.

Let us begin by assuming that some unidentified non-genomic autocatalytic metabolic system (for example, one based on the reverse citric acid cycle), rather than a oligomeric genomic system, did emerge prebiotically, and let us also assume (despite the lack of theoretical or experimental evidence) that such a system would be capable of undergoing Darwinian-type evolution. However, even accepting those far-reaching assumptions it is difficult to see how a prebiotic ***non-genomic*** system would have been able to undergo a series of incremental changes that would have lead to a structurally distinct ***genomic*** system, ***while able to maintain its replicative capability during the transformation***. To see why let us turn to evolutionary theory for useful insights.

Consider Maynard Smith's classic model of protein evolution (Maynard Smith 1970). In his model Maynard Smith made clear that the unitary mutation steps in the amino acid sequence during the evolutionary transformation of protein structure ***cannot pass through non-functional intermediates***. Using that same kind of reasoning leads us to conclude that the transformation of a non-genomic replicating system into a genomic one would have required that each and every step in that transformation also pass through functional intermediates—functional here signifying the possession of a replicative capability—and it is far from clear that such a condition could be satisfied. Just how would the transformation of an autocatalytic metabolic cycle—for example, the reverse citric acid cycle—into a

structurally very different oligomeric sequence-based replicator take place incrementally while maintaining a replicative capability at each and every step of the transformation? The replicative modalities, being quite different, would seem to preclude a smooth transformation in that case. It would be akin to the incremental conversion of a gasoline powered car into an electrically powered one while maintaining the functional capability of the engine at each and every stage! Accordingly, our proposal for a continuous emergence-evolution process, when considered together with a Darwinian model based on the centrality of a genomic oligomer system, reaffirms, we believe, the pre-eminence of a genomic system in the prebiotic phase as well. The suggestion of a peptide-nucleic acid (PNA) oligomer as a possible carrier of earlier genetic information exemplifies this way of thinking (Nielsen 1993), though definitive evidence for any particular pre-RNA entity has yet to be established.

In the context of the Origin of Life debate, Eschenmoser (1994) has expressed the view that chemical theories are significant "if, and only if, they lead to experiments which extend chemical knowledge". In that regard it is important to note that the metabolic vs. genetic replicator dichotomy has an immediate consequence with regard the minimal proto-cell project (Szostak et al. 2001). Both the process of life's emergence and the proto-cell project involve the transformation of inanimate matter into a simple living system. Accordingly, understanding the principles by which Nature chose to undertake this transformation will likely have direct implications on any synthetic attempts to attain that same goal. Biomimetic chemistry par excellence! Thus we conclude by saying that the origin of life field is more than just an intriguing but speculative, academic exercise, as some might suggest. Its deliberations and conclusions, once firmly based, will necessarily impact on the many challenges that still await us at the problematic biology-chemistry interface.

References

Eigen M (1992) Steps toward life: a perspective on evolution. Oxford U.P, Oxford

Eschenmoser A (1994) Chemistry of potentially prebiological natural products. Orig Life Evol Biosph 24:389–423

Eschenmoser A (2007) Question 1: Commentary referring to the statement "The origin of life can be traced back to the origin of kinetic control" and the question "Do you agree with this statement?" Orig Life Evol Biosph 37:309–314

Gesteland RF, Cech TR, Atkins JF eds (1999) The RNA world. Cold Spring Harbor Laboratory Press, NY

Joyce GF (2002) The antiquity of RNA-based evolution. Nature 418:214–221

Kauffman SA (2000) Investigations. Oxford U.P., Oxford

Maynard Smith J (1970) Natural selection and the concept of a protein space. Nature 225:563–564

Mills DR, Peterson RL, Spiegelman S (1967) An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. Proc Nat Acad Sci USA 54:919–927

Morowitz H, Kostelnik JD, Yang J, Cody GD (2000) The origin of intermediary metabolism. Proc Nat Acad Sci USA 97:7704–7708

Nielsen PE (1993) Peptide nucleic acid (PNA): A model structure for the primordial genetic material? Orig Life Evol Biosph 23:323–327

Orgel LE (1998) The origin of life—a review of facts and speculations. Trends Biochem Science 23:491–495

Pross A (2009) Seeking the chemical roots of Darwinism: Bridging between chemistry and biology. Chem Eur J 15:8374–8381

Pross A (2005) Stability in chemistry and biology: Life as a kinetic state of matter. Pure Appl Chem 77:1905–1921
Pross A (2004) Extending the concept of kinetic stability: toward a paradigm for life. J Phys Org Chem 17:312–316
Segre D, Ben-Eli D, Lancet D (2000) Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Nat Acad Sci USA 97: 4112–4117
Shapiro R (2006) Small molecule interactions were central to the origin of life. Quart Rev Biol 81:105–125
Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390
Wachtershauser G (1988) Before enzymes and templates—theory of surface metabolism. Microbiol Rev 52:452–484

E-mail: addy.pross@gmail.com

Liquid Crystal Self-Assembly of Nucleic Acids: A New Pathway for the Prebiotic Synthesis of RNA

Giuliano Zanchetta[1], Tommaso Bellini[1], Michi Nakata[2], Noel A. Clark[2]

[1]Dip. Chimica, Biochimica e Biotecnologie per la Medicina, University of Milano, Italy; [2]Dept. of Physics and Liquid Crystal Materials Research Center, University of Colorado, USA

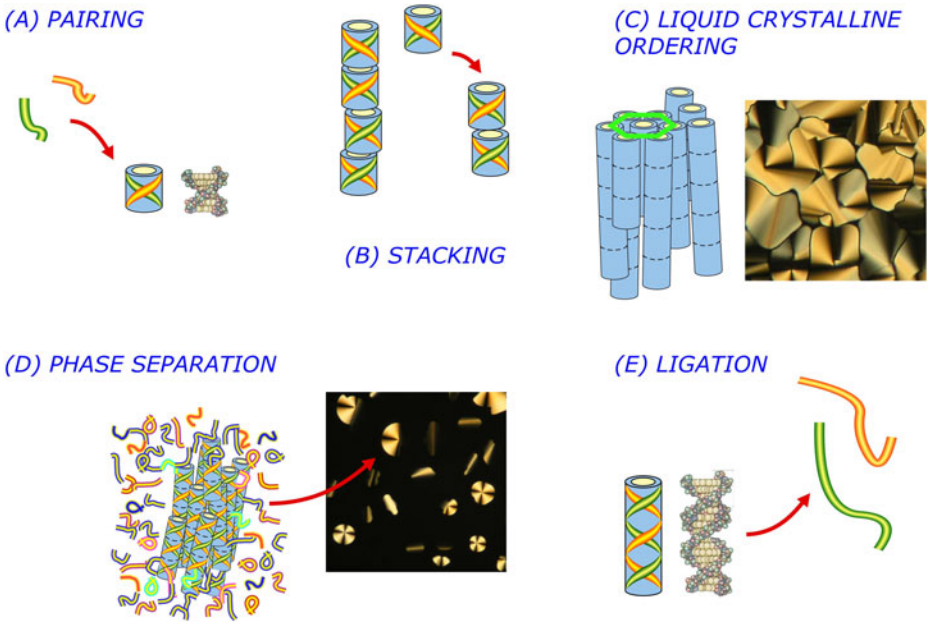
Keywords DNA, RNA • Nucleic Acids • Stacking • Prebiotic • Ribozyme • Ligation • Liquid Crystals

We have recently observed that concentrated aqueous solutions of complementary DNA and RNA oligomers, as short as 6 bp, exhibit chiral nematic and columnar lyotropic liquid crystalline (LC) phases. Structural characterization shows that LC phases are produced by the end-to-end stacking of the duplex oligomers into polydisperse linear aggregates, which are then able to orientationally order (Nakata 2007; Zanchetta 2008a). Furthermore, when only a small fraction of the sequences is complementary, the duplex-forming oligomers segregate from the unpaired strands condensing into LC droplets, thereby maximally concentrating their terminals and holding them in the 3′–5′ configuration favorable to ligation (Zanchetta 2008b). Finally, spontaneous phase separation and liquid crystallization are also found to select complementary and partially complementary strands in concentrated pools of random sequences.

We argue that this set of observations (sketched in the figure) has implications on the self-assembly of nucleic acids and on the prebiotic emergence of RNA strands long enough to sustain the enzymatic activity required by the “RNA world” scenario. Indeed, in the described LC condensation, complementarity promotes concentration via the intermediary of LC order, which provides a natural template for elongation.

Under periodic variation of temperature and/or concentration and in an appropriate chemical environment—conditions possibly met in tidal pools and currently investigated in our labs—, a self-catalytic process may be established for the selective growth of extended complementary strands, since at every ligation longer filaments are formed, that more easily fit within the LC phases and can act as templates for further elongation.

Provided that oligomeric building blocks are available, the described process may act as a proper self-replication mechanism with exponential growth, "selecting" complementary strands but also accommodating some pairing mismatches and thus allowing for copy errors and sequence variation, a prerequisite for the emergence of ribozymes.



References

Nakata M, Zanchetta G, Chapman BD, Jones CD, Cross JO, Pindak R, Bellini T, Clark NA (2007) End-to-end stacking and liquid crystal condensation of 6 to 20 base pair DNA duplexes. Science 318:1276–1279

Zanchetta G, Bellini T, Nakata M, Clark NA (2008a) Physical polymerization and liquid crystallization of RNA oligomers. J Am Chem Soc 130:12864–12865

Zanchetta G, Nakata M, Buscaglia M, Bellini T, Clark NA (2008b) Phase separation and liquid crystallization of complementary sequences in mixtures of nanoDNA oligomers. Proc Natl Acad Sci USA 105:1111–1117

E-mail: giuliano.zanchetta@unimi.it

On Question 6: Plausibility of the 'RNA-world'

Alan Schwartz

Radboud University, Nijmegen, The Netherlands

Keywords RNA World • Chemical Replication • Residue-by-residue Replication

The concept of an RNA world arose out of the series of discoveries which clarified the biological mechanism of inheritance. Very quickly after Watson and Crick, it became increasingly more obvious that DNA and RNA were not only at the heart of the genetic system, but that they must have been involved at the very beginning. With the discoveries related to the catalytic activities of RNA, the theory of an RNA world took final form. This development represented a major breakthrough in thinking about the origin of life (Orgel 2004 and references therein). It can be argued that it is still the only convincing starting point for understanding how biology might have begun. It is important to realize that the hypothesis has been supported by much subsequent work in biochemistry, and that it is *biological evidence*—rather than structural inference or experiments on RNA self-replication—that now provides the strongest argument for the validity of the model (Joyce 2002).

Premise 6 is therefore erroneous in that it assumes that RNA must have been formed prebiotically. In fact, the model tells us little if anything about prebiological chemistry.

The core of the model is the recognition that any precursor molecule (be it nucleic acid, protein, or anything else) must be capable of "residue-by-residue replication" (Orgel 1968). Although chemical replication of carefully designed polypeptides has been demonstrated (Lee et al 1996) this phenomenon has not yet been shown to represent the same level of information transfer as is known with RNA. Chemists have, however, rather convincingly demonstrated that more than one chemical system of information replication would have been capable of chemically driven replication. A number of structural analogs of RNA or DNA have shown interesting properties in this regard, some of which are quite far removed form RNA itself in structure. Even the base-pairing units (the purines and pyrimidines) can be replaced by other molecules which are capable of forming specific complexes between chains in a manner similar to RNA. A question which is not quite so easy to answer, however, is what conditions of concentration and homogeneity would be required to make this idea work on a prebiological earth. Experiments need to be performed to test if the kind of selectivity which has been demonstrated by Bolli et al (1997) for p-RNA in model laboratory experiments, would also apply to reactions carried out with *mixtures* of sequences. It is probable that some type of protocell would have been necessary to permit this level of selection to occur under prebiological conditions.

Alternative hypotheses such as the idea of self-replicating metabolic cycles have been highly touted as models, but no body of data supports the concept, however attractive it may be (Orgel 2008).

References

Bolli M, Micural R, Eschenmoser A (1997) Pyranosyl-RNA: chiroselective self-assembly of base sequences by ligative oligomerization of tetranucleotide-2′,3′-cyclophosphates (with a commentary concerning the origin of biomolecular homochirality). Chem Biol 4:309–320

Joyce GF (2002) The antiquity of RNA-based evolution. Nature 418:214–221

Lee DH, Granja, JA, Severin, MK, Ghadiri, R (1996) A self-replicating peptide. Nature 382:525–528

Orgel LE (1968) Evolution of the genetic apparatus. J Mol Biol 38:381–393

Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. Crit Rev Biochem Mol Biol 39:99–123

E-mail: alan@sci.ru.nl

Workshop OQOL'09

Extended Abstracts for the Following Selected Question

- **Minimal (proto-)cellular world?**

- **Minimal (proto-)cellular world? (a)**

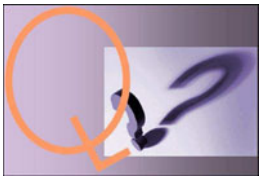
Premise. The simplest cells on Earth contain at least 500–600 genes, and more generally a few thousand. This elicits the question, whether this high complexity is really necessary for cellular life, also in view of the fact that early cells, conceivably, could not have been so complex. Until now, however, the construction of chemical synthetic cells has not been successful, and the attempts to make DNA/Protein "minimal cells" with extant genes and enzymes are still based on systems with approximately a hundred genes. In other words, we are still missing the view of the early protocells—the primitive structures from which modern cells may have arisen.

> **The question**. Do you see a way around the conundrum, that a living cell has to contain several dozens of independent specific macromolecular species and that, nevertheless, this complexity is not reasonably possible in prebiotic times? And/or: how do you envisage the structure of the simplest, early cells?

- **Minimal (proto-)cellular world? (b)**

Premise. The main building blocks of membranes in present-day prokaryotes are rather different from one another: in bacteria (like in eukaryotes) phospholipids are made of fatty acids, linked to the glycerol group (G3P) by ester bonds, whereas the phospholipids of archeabacteria are isoprenoid derivatives linked to glycerol (the stereoisomer, G1P) through ether bonds. And consider the extremely important role of hopanoids and steroids in modern bacterial and eukaryotic membranes.

> **The question**. Do you think that these radical molecular differences show that the issue of compartments was not relevant until late stages in the origin of life? Or do you consider that compartmentalization was still an early landmark, phospholipid diversity being easily explained as a later evolutionary adaptation to extreme environments, for instance?



The Primary Characteristic of Early Cells is very Likely to be the Cell Cycle

Saša Svetina

Institute of Biophysics, University of Ljubljana and Jožef Stefan Institute, Slovenia

Keywords Vesicle Growth • Vesicle Division • Self-reproduction • Cell Cycle • Variability

Cells are individual structured entities that occupy a certain amount of space and are able to grow, replicate and evolve by improving their fitness to the environment. Any answer to the question "How do you envisage the structure of the simplest, early cells?" depends on the answer to "Which characteristic properties of the life process did these cells have, or have to have?" One such property is certainly the maintenance of balance between cell growth and division. This and certain other aspects of the cell cycle behavior discussed below can also be ascribed to vesicles, which supports the notion of vesicular origin of early cells.

Vesicles resemble cells in that they also compartmentalize space. In vesicles the internal aqueous solution is separated from the aqueous environment by a thin flexible membrane. Vesicle membranes can be made of a variety of amphiphilic molecules, including lipids, and the idea about lipid involvement in the emergence of cellular life has been suggested on grounds of the probable abundance of lipids in the prebiotic environment and of their capacity to aggregate spontaneously into micelles or vesicles (Walde 2006). These entities can divide and, in doing so, transmit compositional information to their daughters (Segré et al. 2001). However, the hypothesis about the vesicular origin of cellular life needs to be supplemented by a plausible evolutionary selection mechanism that is not based on gene mutations.

Such a selection mechanism could be based on the process of vesicle self-reproduction. The capacity of growing vesicles to self-reproduce has been proven experimentally (Berklaz et al. 2001). The phenomenon of vesicle self-reproduction was analyzed theoretically on the basis of the requirement to match the growth of the membrane area and vesicle volume with the concomitant transformation of vesicle shape from a single sphere into two spheres connected by a narrow neck (Božič and Svetina 2004; 2007). By attaining this shape, the vesicle is able to split into two daughter vesicles by breaking of the neck. It has been demonstrated that a growing vesicle can reproduce in the described manner only under conditions that combine vesicle growth parameters with the parameters of membrane mechanics. For instance, by analyzing a simple prototype model in which a vesicle is suspended in plain water solution containing only molecules that can be incorporated into its membrane, it was shown that this vesicle can self-reproduce only under the condition that depends on the product $T_d L_p k_c C_0^4$ (= η) where T_d is the vesicle doubling time (related to the growth rate of the membrane area), L_p the membrane hydraulic permeability, and k_c and C_0 the membrane bending constant and spontaneous curvature (Božič and Svetina 2004). If the values of these parameters are such that $\eta<\eta_{cr}$ (where the critical value of η was obtained numerically to be $\eta_{cr}=1.85$), a vesicle would grow into a shape with indistinct characteristics. For $\eta=\eta_{cr}$ the vesicle would reach the shape of two equal spheres connected by a narrow neck (Fig. 1), and for $\eta>\eta_{cr}$ the shape reached would be a combination of two connected spheres of different radii. The generalization of the prototype model that included membrane permeant solvent showed qualitatively analogous behavior, but with a more involved set of conditions for vesicle self-reproduction (Božič and Svetina 2007).

It was a significant realization that the conditions obtained by the described models relate the parameters that depend on environmental conditions such as growth rates, and on vesicle material parameters such as the mechanical properties of their membranes. These conditions can therefore serve as the selectivity criteria for membrane composition. By assuming the selectivity property to be the doubling time for vesicle self-reproduction, one can anticipate the evolution of vesicle systems on the basis of competition between vesicle populations. Vesicle population with an altered membrane composition that can satisfy the criteria but with shorter doubling times would outnumber a vesicle population with the previous membrane composition. Vesicle self-reproduction can thus be considered as the evolutionary process that involves the elements of the Darwinian selection mechanism and can be in a transparent manner also related to the laws of physics and chemistry.

The above conclusion suggests that the early cellular life could have begun by the process of vesicle self-reproduction. Further support for such hypothesis is sought in some generic properties of self-reproducing systems that are so general as to have persisted throughout the process of evolution. In this sense it is of interest to look if the process of vesicle self-reproduction and the cell cycle behavior of contemporary cells share some common features.

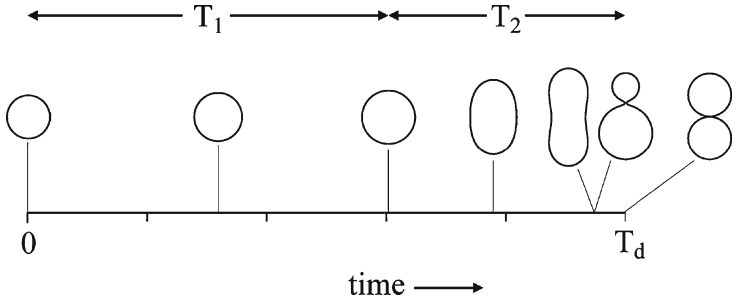


Fig. 1 The axial cross-sections of vesicles predicted by a simple prototype model of the process of vesicle self-reproduction (Božič and Svetina 2004) are presented at the indicated times during its course. T_d is the doubling time. T_1 denotes the duration of vesicle spherical growth and T_2 the duration of its nonspherical growth.

The property pertaining to all self-reproducing systems is their balance between growth and division. The life span of cells can be often divided into their growth phase and the division phase, and in some unicellular organisms like fission yeast this division is extremely sharp (Mitchison and Nurse 1985). In general the rate of cell growth depends to a large extent on the environmental conditions, whereas the act of cell division is more autonomous and primarily dependent on the physical and chemical intrinsic properties of the cell. How cells coordinate growth and division was (Jorgensen and Tyers 2004) and is still (Sawin 2009; Edgar and Kim 2009) a matter of intense research. The dependency of the cell cycle on growth is thought to be established by size requirements for major cell cycle transitions. For instance, in the budding yeast coordination between cell growth and the cell cycle occurs at Start, a short interval in the late G1, i.e. the first phase of the cell cycle during which the yeast commits to division. The Start can be considered as a critical size threshold enforcing a minimal cell size. Different phases can also be assigned to the process of vesicle self-reproduction. The vesicle first grows as a sphere and switches into a phase of non-spherical growth on reaching a certain critical size (Fig. 1) (Božič and Svetina 2004, 2007). The first of these phases can be considered as the growth phase because it is essentially governed by the rate of membrane growth and membrane hydraulic permeability. The transition into non-spherical growth and subsequent shape transformations constitute the division phase the duration of which subtly depends on the mechanical properties of the membrane and thus more on system's intrinsic properties. In vesicles it is thus the conditions for their self-reproduction that serve to control their size. The size control is implemented by the selection of vesicle properties that affect the process of its division. Treated vesicle systems also taught us about the increase of the complexity of the self-reproduction criteria as the system becomes more complex (Božič and Svetina 2007). By assuming that in their early evolution cells followed the same principles, the criteria for their ability to replicate must have become

extremely complex. It is probably why the insights into the cell size control mechanisms have until recently been relatively sparse (Umen 2005).

In many cellular systems the inherent feature of the interplay between cell growth and division is also the variability of cell generation times (Di Talia et al. 2007; Tsur et al. 2009). In general, the G1 phase of the cell cycle is essentially more variable than cell cycle phases that follow cell's commitment to division. This variability could be ascribed to molecular noise in gene expression but it was shown at least for the budding yeast that it is also the consequence of different sizes of daughter cells and the existence of the size control (Di Talia et al. 2007). The variability of vesicle doubling time occurs in the case of $\eta > \eta_{cr}$ where daughter vesicles attain different initial sizes. The smaller daughter vesicle must grow longer than the larger to reach (in the prototype model actually the same) critical size. The duration of the phase of the non-spherical growth (T_2 in Fig. 1) is the same for both daughter vesicles (Božič and Svetina 2004). The variability of the vesicle generation time is thus the consequence of the variability of the spherical part of its growth which can be compared to the larger variability of the G1 phase of the cell cycle.

The vesicle version of the origin of cellular life implies that after the process of vesicle self-reproduction was properly initiated, it had to be continued without interruption. This notion also provides a possible answer to Question 9 (Life as unity or confederacy?). As already suggested (Svetina 2007), in view of the so called vesicle world the vesicular systems could have increased their metabolic and organizational complexity on the basis of diversification of vesicle properties and fusion of vesicles carrying different features. However, it has to be kept in mind that the correct structural and functional inheritance of contemporary cells and their ability to evolve are based on genes. In view of the vesicle world the cell genetic structure can be considered as one of the evolutionary improvements. Therefore, in order to confirm the idea about the vesicle world, the major problem to be solved will be to understand how the process of transmission of information could have been switched from the compositional to the contemporary genetic one.

References

Berklaz N, Müller N, Walde P, Luisi PL (2001) Growth and transformation of vesicles studied by ferritin labeling and cryotransmission electron microscopy. J Phys Chem B 107:1056–1064

Božič B, Svetina S (2004) A relationship between membrane properties forms the basis of a selectivity mechanism for vesicle self-reproduction. Eur Biophys J 33:565–571

Božič B, Svetina S (2007) Vesicle self-reproduction: The involvement of membrane hydraulic and solute permeabilities. Eur Phys J E 24:79–90

Di Talia S, Skotheim JM, Bean JM, Siggia ED, Cross FR (2007) The effect of molecular noise and size control on variability in the budding yeast cell cycle. Nature 448:947–951

Edgar BA, Kim KJ (2009) Sizing up cells. Science 325:158–159

Jorgensen P, Tyers M (2004) How cells coordinate growth and division. Curr Biol 14: R1014–R1027

Mitchison JM, Nurse P (1985) Growth in cell length in the fission yeast Schizosaccharomyces pombe. J Cell Sci 75:357–376

Sawin KE (2009) Cell division brought down to size. Nature 459:782–783

Segré D, Ben-Eli D, Deamer DW, Lancet D (2001) The lipid world. Orig Life Evol Biosph 31:119–145

Svetina S (2007) The vesicle world: The emergence of cellular life can be related to properties specific to vesicles. Orig Life Evol Biosph 37:445–448

Tzur A, Kafri R, LeBleu VS, Lahav G, Kirschner MW (2009) Cell growth and size homeostasis in proliferating animal cells. Science 325:167–171
Umen JG (2005) The elusive sizer. Curr Opin Cell Biol 17:435–441
Walde P (2006) Surfactant assemblies and their various possible roles for the origin(s) of life. Orig Life Evol Biosph 36:109–150

E-mail: sasa.svetina@mf.uni-lj.si

Constructive Approach to Proto-cellular Life

Tetsuya Yomo

Department of Bioinformatics Engineering, Osaka University, Japan

Keywords Artificial Cell • Cell-free Translation • RNA Replication • Self-replication • Qbeta Replicase • Liposome

We have been working on the experimental synthesis of the lipid vesicle in which the genetic molecule (RNA) is self-replicated by its encoding protein (RNA-dependent RNA polymerase). As the experimental setup for the *in liposome* RNA-protein self-replication network is not close enough to the pre-biotic soup, we cannot answer to the question as to how the primordial life appeared at the beginning. However, carefully avoiding conclusions that strongly depend on which kinds of molecules were used in the experiments, we will discuss basic questions on the requirement of small compartments and the upper limits of the composition complexity for the origin of life.

Not in small compartment but in a bulk solution, the simplest self-replication system composing a genetic molecule an encoded protein, each of which catalyzes the synthesis of each other, the replication will accelerate if the concentration of either of the components increases due to the nature of the mutually catalytic network. Resultantly, the concentration of all the components could have evolved to turn higher and higher. But when it comes to living compartments, e.g., present cells, DNA molecule is a few in copy number while proteins are many. If there were many copies of DNA in a cell, the genetic information would lose the evolvability, because the effect of many mutations independently occurred on many DNA molecules is averaged-out in the single cell so that the phenotypic diversity among the cells would turn out to be too small for Darwinian selection. Such hypothesis was demonstrated here by the selection experiment on the lipid vesicle encapsulating the transcription and translation processes. DNA must be as few as possible in number, in addition to other requirements to be genetic molecule; the heritability and combinatorial power of encoding (Sunami et al. 2006).

This small-number rule for genetic molecule per cell limits the compartment size by the diameter of 10 μm. Supposed that a single gene produces 1,000 molecules of a protein catalyst in compartments of 10 μm diameter, the concentration of the catalyst will be about 1 nM, which is approximately equal order of dissociation constants for natural DNA binding proteins of high affinity. If the compartment size were larger than 10 nm diameter, the concentration would be too small for the catalyst to re-bind to the DNA or RNA for the self-replication reaction.

How many components are needed to run the self-replication in lipid vesicles? The self-replicating system was assembled using one template RNA encoding an RNA-dependent

RNA polymerase which is translated by in vitro translation system reconstructed from purified translation factors and replicates the original template RNA. We have successively run this RNA-protein self-replicating system *in liposome* only with 144 bio-polymers. Although this number may vary slightly depending on the experimental setup, the order was close to the estimated minimal number of the essential genes presently reported. This experimentally demonstrated composition complexity is too high for the origin of life but shows a level of complexity that the origin of life went through at a certain time point in the developing process to present cells (Kita et al. 2008).

References

Sunami T, et al (2006) Femtoliter compartment in liposomes for in vitro selection of proteins. Anal Biochem 357:128–136

Kita H, et al (2008) Replication of genetic information with self-encoded replicase in liposomes. ChemBioChem 9:2403–2410

E-mail: yomo@bio.eng.osaka-u.ac.jp

Conceptual bases for the emergence of early protocells

Gabriel Piedrafita[1], Felix Olasagasti[2], Kepa Ruiz-Mirazo[3], Juli Peretó[4], Álvaro Moreno[3], Francisco Montero[1], Federico Morán[1]

[1]Dpto. de Bioquímica y Biología Molecular I, Facultad de CC. Químicas, Universidad Complutense de Madrid, 28040 Madrid, Spain; [2]Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, USA; [3]Dpto. de Lógica y Filosofía de la Ciencia Universidad del País Vasco UPV/EHU, Apto de correos 1249, 20080 Donostia/San Sebastian, Spain; [4]Institut Cavanilles de Biodiversitat i Biologia Evolutiva y Dpto. de Bioquímica i Biologia Molecular, Facultat de Ciències Biològiques, Universitat de València, València, Spain

Keywords Autopoiesis • Self-production • Thermodynamic Coherence • Self-assembly

Life as we know it today exhibits both a *self-maintaining* and a *self-reproducing* organization. In the chemical domain, self-maintenance involves self-production or 'autopoiesis' (Varela et al., 1974), i.e., a network of components and transformation processes (chemical reactions) that achieves a cyclic, recurrent dynamics, which is implemented in thermodynamically open but operationally or functionally "closed" systems. Through the continuous generation and regeneration of their own components, the organization of these systems remains the same, despite the structural-physical changes taking place in them. However, self-producing systems may develop the potential to increase their complexity, provided that they become also self-*re*productive and start an evolutionary process that will allow them to explore other forms of organization, by incorporating new types of components, processes, and associated functions.

Although some currents of thought in the fields of theoretical biology and origins of life do not consider the compartment as a crucial requirement to achieve this type of cyclic, self-productive organization (Rosen 1991; Kauffman 1993; Wächtershäuser 1990) there are good

reasons to consider that the capacity of a system to create and maintain its own boundaries is part and parcel of that process (Varela et al. 1974; Ganti 1975; Ruiz-Mirazo et al. 2004) and, furthermore, it should be an early landmark in prebiotic evolution (Oparin 1961; Morowitz 1992; Deamer 1997). In any case, two main features should have characterized early protocells: (1) be simple enough to appear easily in the chemical conditions of the primitive Earth; (2) capacity to allow further prebiotic evolution toward more complex systems, acquiring new properties. In other words, the system dynamics must provide a framework that gives certain independence or autonomy from the environment, at the same time as it is able to preserve itself and to generate new components of increasing structural complexity.

Focusing on the plausibility of such self-producing proto-cellular systems and their evolutionary development in different prebiotic scenarios, some general conditions should be addressed:

a) Materiality. Prebiotic systems were built from common materials and molecular compounds on the early Earth. Extant cells are a form of physical-chemical organization of matter. Metabolic reactions represent a particular subset of the whole of realizable chemical reactions in which catalysts act as constraints modifying the dynamics of inter-conversion processes among the different components and building-blocks of the system.
b) Recursive self-maintenance. As already mentioned, the condition of recursivity in terms of organizational closure is needed for self-maintenance (Cornish-Bowden & Cárdenas 2007).
c) Thermodynamic coherence. Since the system must be open, i.e., capable to exchange energy and/or matter with the environment, the processes within it must be energetically and mechanistically coupled. There must be a subsystem of energy "currencies", as it happens in present cells (Skulachev 1992), to guarantee that the flow of matter and energy across a membrane is adequately coupled with the internal proto-metabolic reaction network, responsible for the production of that very membrane (Ruiz-Mirazo & Moreno 2004).
d) Stoichiometric coherence. Stoichiometric balance is needed, as a direct consequence of the mass conservation law acting on the chemical reaction network with extension to open systems (Montero et al. 2008).
e) Self-reproduction. The system must have a specific way to reproduce itself. This property would be a consequence of a particular way of realisation of autopoiesis in which productive processes outweigh or prevail over decay processes and there is net growth in the system. Eventually this may include the possibility to generate differences on network progeny efficiency and competition, allowing the increase of complexity in some sort of pre-Darwinian evolution.
f) Increase of structural and functional complexity. Structural complexity could be achieved by subsequent cycles of system growth and reproduction, which will favour the incorporation of new components in the network. Increases in functional complexity would be related to new types of constraints that enter a mutually reinforcing loop, like the combined action of a membrane and a set of catalysts providing kinetic confinement (Moreno & Ruiz-Mirazo 2009). Finally, all this generated complexity must be kept or fixed in the system (by several possible mechanisms, such as redundance, feed-back, or buffering) and in its offspring (template replication, various forms of heredity…).

How to envisage the structure and nature of the first protocells?

As a direct response to the question raised during the workshop, we consider that before a minimal cell with a proper genome was developed (i.e., a minimal living cell), other types of protocellular systems that by and large fulfil the previous requirements had to appear. For

instance, a theoretical model of an energetically self-maintaining system has been recently proposed (Olasagasti et al. 2007; Montero et al. 2007) satisfying some of the aforementioned requirements (Figure 1A). This model illustrates how a protocell could manage to satisfy its own energetic or thermodynamic needs with a relatively autonomous way of operating and harvesting some external supply of energy. It includes two membrane components—Tm and Em—that are synthesized within the system and act respectively to generate a chemiosmotic gradient and profit it to produce an energy-rich compound—A—which is itself the driving force for synthesizing all the system-distinctive molecules, including Tm and Em too. In this sense, the whole system acts as a functional entity where every reaction appears conveniently coupled (both kinetically and thermodynamically) and the active role of the membrane permits the flow control of matter-energy that is essential for the maintenance of any self-constructive organization in a changing environment. Nevertheless, this model is quite abstract, and this hinders its own understanding as a model simple enough to describe a plausible early protocell, maybe because it implies transmembrane molecules acting as energy transducers, which nowadays evoke rather complex molecules, i.e. proton pumps or ATPase-like motors.

In an attempt to advance towards more realistic protocell designs, more explicit models have been proposed, taking into account the nature and special properties of some of its basic components, such as the capacity of some simple lipids to self-assemble into close bilayers. The model in Figure 1B tries to explain the transition to self-maintaining autopoietic organizations, starting from bare self-assembled vesicles (Piedrafita et al. 2009). A complex lipid molecule "L" starts to be produced from a simple naturally occurring lipid "l", making use of some high-free-energy precursors (X and Y) that cross the vesicle membrane, which is entrapping a catalyst A_i. In this case the main problem is that the maintenance of the self-constructive organization depends on the external availability of the precursors, so this heterotrophic-type of system results very dependent on the environmental conditions. Furthermore, the lack of a proper control of matter-energy flow can lead to osmotic imbalance and collapse.

The matter-energy flow control problem could be avoided by considering the acquisition of a higher degree of "autonomy", through the interplay between the boundary and some internally produced component, which is inserted spontaneously in the membrane modifying its properties, e.g. its permeability (Ruiz-Mirazo & Moreno 2004). In the Figure 1C, the active role of the compartment is explained because of the appearance of a polymer P_n capable of changing the membrane permeability and thus, regulating the relationship with the environment (Ruiz-Mirazo & Mavelli 2008). In this way, the protocell benefits from its own self-constructive dynamics, making itself more independent from the external constraints and guaranteeing a more robust maintenance.

The cornerstone is still to find a particular set of available compounds that could generate some kind of minimal but evolutionary interesting self-maintaining cellular organization. The above theoretical models try to provide a better explanation of how early protocells could have developed before achieving highly complex tasks, such as the synthesis of informational genetic molecules. Nevertheless, the physical-material implementation of those protocellular reaction network schemes is not easy to achieve. Future efforts should be focused on reducing the gap between in silico and in vitro approaches, solving the characteristic level of abstraction of theoretical models with a proper combination with real lab experiments, in order to get better insights on the nature and dynamic properties of the early protocells. In this sense, it could be appealing to study the possible role of short peptides in the implementation of functionalized compartments, and explore the lipid-peptide scenario as a starting point for basic autonomy, with potential to be developed in subsequent evolutionary transitions.

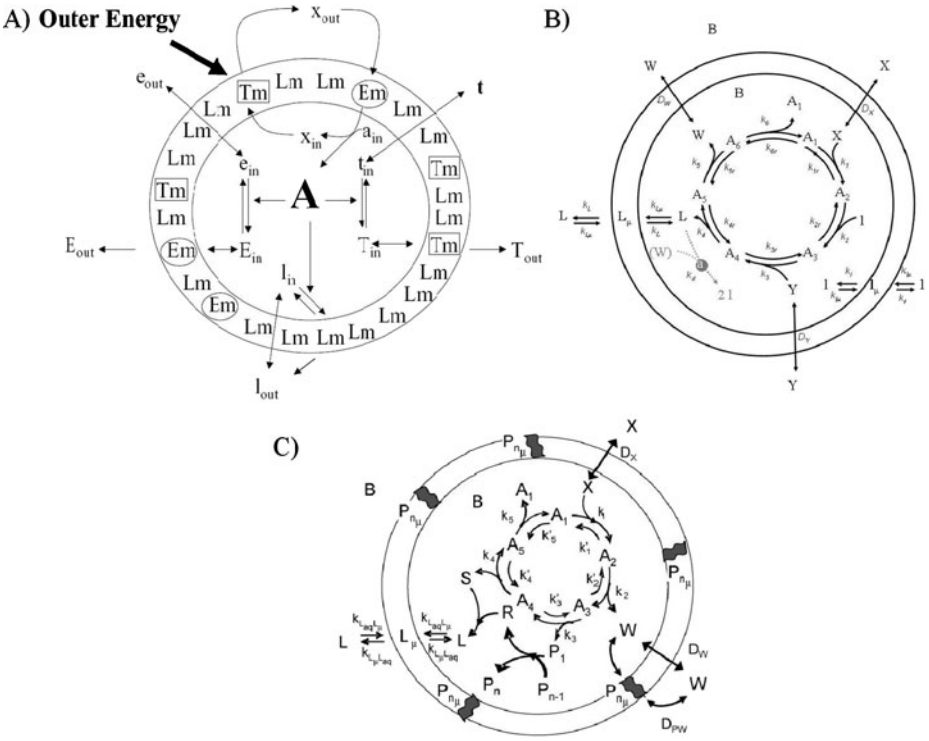


Fig. 1 Different protocell models developed by the authors: Olasagasti et al. 2007 (A); Piedrafita et al. 2009 (B); Ruiz-Mirazo and Mavelli 2008 (C).

References

Cornish-Bowden A, Cárdenas ML (2007) Organizational invariance in (M,R) systems. Chem Biodivers 4:2396–2406

Deamer DW (1997) The first living systems: a bioenergetic perspective. Microbiol Mol Biol Rev 61:239–261

Ganti T (1975) Organization of chemical reactions into dividing and metabolizing units: the chemotons. BioSystems 7:15–21

Kauffman S (1993) The origins of order: Self-organization and selection in evolution. Oxford Univ. Press, Oxford

Montero F et al (2008) Stoichiometric analysis of self-maintaining metabolisms. J Theor Biol 252(3):427–432

Moreno A, Ruiz-Mirazo K (2009) The problem of the emergence of functional diversity in prebiotic evolution. Biology and Philosophy 24:585–605

Olasagasti F et al (2007) Energetically plausible model of a self-maintaining protocellular system. Bull Math Biol 69(4):1423–1445

Oparin AI (1961) Life: its nature, origin and development. Oliver and Boyd, Edinburgh

Piedrafita G et al (2009) On the transition from prebiotic to proto-biological membranes: from 'self-assembly' to 'self-production'. Springer LNCS. In press
Rosen R (1991) Life itself: a comprehensive inquiry into the nature, origin and fabrication of life. Columbia Univ. Press, New York
Ruiz-Mirazo K, Moreno A (2004) Basic autonomy as a fundamental step in the synthesis of Life. Artif Life J 10(3):235–260
Ruiz-Mirazo K, Mavelli F (2008) On the way towards 'basic autonomous agents': stochastic simulations of minimal lipid-peptide cells. BioSystems 91(2):374–387
Skulachev VP (1992) The laws of cell energetic. Eur J Biochem 208:203–209
Varela F et al (1974) Autopoiesis: the organization of living systems, its characterization and a model. BioSystems 5(4):187–196
Wächtershäuser W (1990) Evolution of the first metabolic cycles. Proc Natl Acad Sci USA 87:200–204

E-mail: fmoran@bio.ucm.es

Early Origin of Phospholipid Membranes but Late Specialisation of Membrane Components

Purificación López-García[1], Juli Peretó[2], David Moreira[1]

[1]Unité d'écologie, Systématique et evolution, Université de Paris-Sud, Orsay, France;
[2]Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de Valencia, Spain

Keywords Cell Membranes • Cenancestor • Fatty Acids • Isoprenoids • Last Universal Common Ancestor • Phospholipids

There are two basic kinds of membrane phospholipids in extant organisms. Bacterial and eukaryal membrane phospholipids are fatty acid esters linked to *sn*-glycerol-3-P (G3P). Archaea possess isoprenoid ethers built on *sn*-glycerol-1-P (G1P). While there are some exceptions to the nature of the lateral chain, fatty acid or isoprenoid, and that of the linkage (ether versus ester), exceptions to the opposite chirality of the two types of phospholipids have never been observed. Since the two key enzymes leading to G1P and G3P, G1P- and G3P- dehydrogenase (G1PDH and G3PDH), are not homologous, they might have originated during the speciation of the two prokaryotic domains, opening the possibility that the last universal common ancestor (cenancestor) lacked a membrane (was acellular) or possessed no phospholipid membranes at all. We have previously shown that G1PDH and G3PDH belong to two separate superfamilies universally distributed, suggesting that members of both superfamilies existed already in the cenancestor. We suggested that the cenancestor was capable of synthesizing phospholipids enzymatically but leading probably to mixtures of lipids based on both glycerolphosphate stereoisomers. We also showed that many archaea possess homologues to nearly all known bacterial genes involved in fatty acid metabolism, showing the potential to synthesise fatty acid phospholipids. This, together with the presence of universally conserved proteins intimately linked to membranes, such as those of a respiratory chain and membrane ATPases used to generate free energy to the expense of a chemiosmotic gradient, argues in favour of a cenancestor already endowed with membrane phospholipids and an earlier origin of the lipidic nature of the cellular compartment.

References

Peretó J, López-García P, Moreira D (2004) Ancestral lipid biosynthesis and early membrane evolution. Trends Biochem Sci 29:469–477

E-mail: david.moreira@u-psud.fr

Phospholipids May Have Co-evolved with Chemiosmosis

Robert Pascal

Institut des Biomolécules Max Mousseron, University of Montpellier 2, France

Keywords Chemiosmosis • Fatty Acids • Glyceraldehyde-1-phosphate • Glyceraldehyde-3-phosphate • Hydrogenases • Phospholipids

According to Darwinian theory, evolution is the result of a combination of contingency and of determinism playing the role of driving force. The early evolution of the membrane has no reason to have escaped this scheme. In this view, the utility of the membrane, even at early stages, constituted a driving force orientating the selection of its components among multiple possibilities. There is increasing evidence that fatty acid are able, alone or in combination with other components (Deamer 1997), of promoting the formation of bilayer membranes and then of compartmentalization. Their participation to the early development of life is then highly likely since compartment formation is a simple mean of avoiding the dispersion of components, and especially macromolecular ones. Fatty acid-made vesicles could have harboured genetic polymers while their permeability was favourable to a spontaneous exchange of metabolites (Mansy et al. 2008) and alkali cations (Chen and Szostak 2004). As the metabolism became more complex new kinds of hydrophobic substances became probably available to build the membrane. But free fatty acids are no longer present as membrane components, an observation that raises the question of which driving force led to discard them. The idea that fatty acids became disadvantageous when chemiosmotic pathways of harvesting energy became available favouring organisms able to maintain a ion gradient allowing then the formation of ATP by coupling ion transfer with the formation of a phosphate anhydride is developed here.

Although lipids and other saturated hydrocarbon derivatives are among the most energetic substrates for aerobic respiration, it is worth to notice that in an anoxic environment these substance are conversely almost devoid of energetic value, which accounts for their long term stability in sediments following their formation through biological or abiotic hydrogenation processes (Hebting et al. 2006). Then, since they are located near the bottom of a potential energy well in a reducing environment, their formation could be the result of the metabolism of precursors having higher degrees of oxidation such as sugars (Weber 2000). Moreover, their participation to membrane stabilization constituted a decisive advantage supporting an early emergence of compartments. Actually, bilayer membrane vesicles are formed from fatty acids when the solution pH is near the apparent pK_A of the membrane-incorporated fatty acids (Hanczyc and Szostak 2004). Vesicles made of these surfactants may have harboured genetic polymers, while their permeability was favourable to a spontaneous exchange of metabolites (Mansy et al. 2008) and alkali cations (Chen and Szostak 2004). But the translocation of ions strongly restricted the possibility of development of stable proton concentration gradients between the two sides. Even the fast translocation of the neutral form of fatty acids (Simard et al.

2008) is by itself able to hinder the formation of pH gradients. On the early Earth, this property was probably amplified by a lower pH induced by a higher CO_2 content in the atmosphere (Martin et al. 2006). Even when present as minor components in phospholipid-based membranes, fatty acids are capable of a similar behavior (Simard et al. 2008). Owing to electrostatic interactions with the anionic environment, fatty acids can display apparent pK_A values of up to 7.5 in a phospholipid membrane, which allows them to translocate at high rates, but the proton flux is limited and unidirectional. In biology, the ability of fatty acids to restrict the development of a proton gradient across a phospholipid membrane by translocating as neutral species (Pick 1987; Kadenbach 2003) has been shown to uncouple the synthesis of ATP through chemiosmosis. In the evolution of life, the ability to take advantage of ion gradients to build ATP very probably appeared much later than compartments since it requires membranes that are impermeable to ions. The achievement of this task required both the selection of surfactants with a low pK_A such as phospholipids remaining in a fully ionized state (Westheimer 1987) and the presence of transporters for substrates and metabolites to compensate for the low permeability of the thus evolved barrier. This scenario is also compatible a two stage process in which Na^+-translocating ATPases (Mulkidjanian et al. 2008) emerged before H^+-translocating ATPase since a spontaneous translocation of alkali metal cations is possible only for membranes with high contents in fatty acids (Chen and Szostak 2004). In this discussion on lipid diversity, it is then proposed that chemiosmosis co-evolved with the development of efficient biochemical pathways to get rid of fatty acids in the membrane.

The scenario required (i) phospholipids-based membranes (ii) the availability of a biochemical machinery capable of confining the fatty acid concentration below the threshold needed for a stable enough proton concentration gradient. Since the availability of a substantial supply of ATP gave a high selective advantage to the corresponding organisms, it constituted a driving force for the selection of biochemical pathways avoiding free fatty acids, which was accompanied by the loss of incompatible ones. The formation of phospholipids from glyceraldehydes-3-phosphate (G3P) in Bacteria and glyceraldehydes-1-phosphate (G1P) in Archaea may be two different strategies of evolution as a consequence of this event. Before the emergence of chemiosmosis the lipid composition of the membrane would have been much less efficiently controlled including different kinds of membrane lipids such as fatty acids and many other amphiphilic molecules including possibly G1P and G3P derivatives. At that early stage, lipids, formed abiotically or as by-products of carbon metabolism, could be used as building blocks for compartmentalization. Then, evolution selected solutions among a large diversity of metabolites, those that allowed chemiosmosis and the delivery of a huge amount of energy produced by membrane ATPases to other metabolic pathways. This scenario, which introduces a further selective constraint (the ability to endure chemiosmosis) at a late stage provides an explanation to the fact that two main solutions have been retained in the living world: fatty acids bound by ester linkages to G3P in bacteria (and Eucarya) and isoprenoid lateral chains that are bound by ether linkages to G1P in archaea. These views are compatible with the hypothesis that the cenancestor synthesized glycerol phosphates in a non-stereoselective way and that hydrogenases able to produce either G1P or G3P evolved later (Peretó et al. 2004), which may be the consequence of a selective process driven by the biochemical utility of these intermediates.

Acknowledgments The author is grateful to P. López-García and D. Moreira for helpful discussions.

References

Chen IA, Szostak JW (2004) Membrane growth can generate a transmembrane pH gradient in fatty acid vesicles. Proc Natl Acad Sci USA 101:7965–7970

Deamer DW (1997) The First Living Systems: a Bioenergetic Perspective. Microbiol Mol Biol Rev 61:239–261
Hanczyc MM, Szostak JW (2004) Replicating vesicles as models of primitive cell growth and division. Curr Opin Chem Biol 8:660–664
Hebting Y, Schaeffer P, Behrens A et al (2006) Biomarker Evidence for a Major Preservation Pathway of Sedimentary Organic Carbon. Science 312:1627–1631.
Kadenbach B (2003) Intrinsic and extrinsic uncoupling of oxidative phosphorylation. Biochim Biophys Acta 1604:77–94
Mansy SS, Schrum JP, Krishnamurthy M et al (2008) Template-directed synthesis of a genetic polymer in a model protocell. Nature 454:122–125
Martin H, Caeys P, Gargaud M et al (2006) Environmental Context. Earth Moon Planets 98:205–245
Mulkidjanian AY, Galperin MY, Makarova KS et al 2008. Evolutionary primacy of sodium bioenergetics. Biol Direct 3:13
Pereto J, Lopez-Garcıa P, Moreira D (2004) Ancestral lipid biosynthesis and early membrane evolution. Trends Biochem Sci 29:469–477
Pick U, Weis M, Rottenberg H (1987) Anomalous Uncoupling of Photophosphorylation by Palmitic Acid and by Gramicidin D. Biochemistry 26:8295–8302
Simard JR, Pillai BK, Hamilton JA (2008) Fatty Acid Flip-Flop in a Model Membrane Is Faster Than Desorption into the Aqueous Phase. Biochemistry 47:9081–9089
Weber AL (2000) Sugars as the optimal biosynthetic carbon substrate of aqueous life throughout the universe. Orig Life Evol Biosph 30:33–43
Westheimer FH (1987) Why nature chose phosphate. Science 235:1173–1178

E-mail: rpascal@univ-montp2.fr

Minimal and Primitive Cells in Origins of Life, Systems Chemistry and Synthetic Biology

Pasquale Stano

Department of Biology, Univ. RomaTre, Rome, Italy

Keywords Minimal Cells • Liposomes • Protocells • Systems Chemistry • Synthetic Biology

One of the many unsolved questions in origin of life research concerns the appearance of primitive cells. Since we do not really know the primitive cells, we generally use this term to indicate a not well defined family of structures that preceeded the fully-fledged living cell. So, we first have to recognize that primitive cells (protocells) may have different degrees of complexity, i. e., from simple compartmentalized molecules to more sophisticated structures; then we need to clarify at which point—along this hypothetical scale of increasing complexity—the property of being "alive" can (or cannot) be recognized. Both aspects are of great relevance: the first because it requires to draw a sort of plausible pathway with several intermediate steps from very primitive molecules to a multimolecular system capable of a coordinate and coherent behavior; the second because it asks the intricate question of defining a minimal set of properties that are necessary and sufficient for being alive. It is evident that defining the structure, the functions and the properties of the *simplest* cell—yet *living*—is the most attractive and challenging goal in this field.

No fossil protocells are however available, and we can first develop a scientific approach from theoretical considerations, but when we have to focus to more concrete model, this enterprise becomes very difficult.

A top-down approach to the minimal cell

We can start our consideration from two well established scientific facts: (1) the simpler/smaller cells that we know—being free-living or parasitic organisms—still have a rather complex biochemical network and contain several hundred of genes, enzymes implementing fundamental cellular functions; (2) comparative genomic analysis points out that a hypothetical minimal genome is composed by 206 genes (Gil et al. 2004), and that the corresponding organism would survive in a highly permissive environment, that provides the cell with all the required compounds that it cannot synthesize. On this basis, we can define a *minimal* living cell, based on modern DNA-RNA-protein machinery, whose structure and functions are derived from a top-down analysis of extant cells. The minimal genome accounts for basic cellular functions such as DNA duplication, protein synthesis, core metabolism, etc. (Gil et al. 2004). Despite the large number of chemicals and processes associated to a minimal cell with 206 genes, it is useful to consider what this minimal cell actually does, forgetting for a moment the details of each molecular process. It is easy to show that—in essence—this hypothetical minimal cell simply produces a set of molecular components (enzymes, ribosomes) required to setting up a series of processes that brings to the production of all its molecular components, including structural components, from some available building blocks.

Said in different way, the definition of the minimal cell functions follows a "circular" logic. In the minimal cell, a minimal set of compounds are structurally and functionally organized in a compartmentalized fashion in order to produce processes, that in turn lead to the synthesis of these compounds, etc., with the net result of keeping "alive" the structural and functional cell organization, at the expenses of changes in the environment. This is, actually, the definition of living cell that is proper of autopoiesis, the elegant conceptualization provided by Humberto Maturana and Francisco Varela in the Seventies (Varela et al. 1974).

It is also worth noting that the minimal cell obtained by the top-down approach is locked onto the figure of 206 genes simply due to the fact that it uses sophisticated, efficient, and specific molecular machines (enzymes, ribosomes) to carry out the required processes. Although it may turn to be very difficult to synthesize a minimal cell with 206 genes in the laboratory, we have to admit that the minimal cell is somehow simpler than all known living cells. In this context, however, simple does not mean primitive.

Primitive cells

A conceptual bridge between the minimal cell (as obtained by top-down analysis) and the primitive cells does indeed exist, and it actually encompasses a family of hypothetical cells, like fully-synthetic or artificial cells (Rasmussen et al. 2004), semi-synthetic cells (Luisi et al. 2006), ribozyme-based cells (Szostak et al. 2001) (see Box 1). The common denominator is the establishment of the above-mentioned ciclic logic of autopoiesis, that applies to every model of living cell, regardless its material implementation.

Now, to derive the structure of primitive cells from minimal DNA-protein cells, we do not need to simply reduce the number of their components, since this number depends contingently on the modern biochemistry. Moreover, since modern DNA, RNA and enzymes coevolved in a mutual dependent way every attempt to isolate one component from its context does not provide useful insights into the nature of primitive cells. In other words, it turns out that a cell provided by the minimal genome has already reached an irriducible complexity. A simplification is possible only acting by a sort of quality reduction (as opposite to quantity reduction), identifying how the core processes of self-producing can be implemented by simpler molecular systems. Clearly, for the continuity principle, the nature of primitive cells must be compatible to what we know about prebiotic chemistry and modern cells.

At the current stage of knowledge, only speculations can help. Suppose first that the main scheme of modern biology is conserved (DNA to RNA to proteins and then functions). In this case, a possible reduction of minimal cell components and processes is possible only in limited way. For example, a less specific polymerase could catalyze the synthesis of both DNA and RNA, or a less specific amino acyl-tRNA synthetase could act on several substrates at the cost of lower fidelity. Other classical examples are "protoribosomes" with only few simple basic proteins, or metabolic enzymes which are compatible with a wider range of substrates. In this way, the number of different macromolecules required to accomplish the minimal cell functions can be indeed reduced, but probably not more than a factor 2.

For further simplifications, we have to radically reconsider the nature of primitive cells, for example by employing only a subset of the 20 amino acids, or thinking about catalysis by short peptides and by ribozymes, and in general by departing from the modern way of encoding functions (linear sequence in nucleobase space → linear sequence in protein space → self-organization in a three dimensional functional structure). The first obvious way to simplify this route is to employ a functional molecule that can also store and transmit its structure, and this would correspond to cells entirely based on ribozymes. An attempt to define primitive cells in term of lipids and ribozymes has been provided by Szostak, Bartel and Luisi in a recent conceptual script (Szostak et al. 2001).

In more general and primitive terms, however, simpler and more basic protocell design should be conceived, and the adoption of a systemic paradigm is certainly advantageous. We have to look for in terms of spatially organized, self-bounded and self-producing autocatalytic reaction network, as required by autopoiesis, that is actually a system theory. How to translate these general statement into real chemical systems? And, moreover, how to identify such a system that ultimately can evolve to life as we know it? It is true, in fact, that despite the numerous theoretical studies on this subject, the corresponding experimental investigation is largely missing.

Can current synthetic biology and systems chemistry approaches help in investigating primitive cells?

In order to answer to the fourth question of the international workshop on the Open Questions about the Origins of Life (San Sebastian, Spain) (on the incompatibility of minimal cell complexity and primitiveness; and consequently on our ignorance about the structure of early cells), a first point for clarification is that we simply do not know and perhaps we will never know the exact historical sequence of the events that leaded to living cells. In my opinion, such historical reconstruction is impossible not only for practical reasons, but also conceptually (since it cannot be proved that things happened in a certain way).

Therefore our efforts should be directed toward the demonstration that the core functions of a primitive cell can be carried out by molecules which are compatible with the most simple chemistry we can imagine, and this is not the modern molecular biology. In order to understand the nature of early cells it is important to look to classic fields as supramolecular chemistry, and fully exploit the mechanisms of self-assembly and self-organization. The newly dubbed field of *systems chemistry* specifically looks at the problem of molecular systems from these viewpoints, also encompassing most of the modern aspects of organocatalysis, autocatalysis and reciprocal (network) catalysis. The construction of minimal cells by this bottom-up approach, which does not explicitally look to the origins of life problem, relies on the above-mentioned bioinspired mechanisms and might give many hints to understand core aspects of compartmentalized reacting systems. One of the most interesting concept of modern chemistry is organocatalysis (Nature Insight 2008). Small molecules can indeed act as catalysts, and this is not surprising, because enzymes are nothing else than large organocatalysts, but their functional groups are kept in proper

spatial arrangement by the rest of the molecule. The catalysis by small molecule and by molecular complexes is probably the key aspect of primitive cells, which could be constituted by such kind of simple molecules.

To be more specific, a recent report shows that a simple dipeptide, such Ser-His, is capable of forming peptide bonds starting from ethyl esters of amino acids and free amine (Gorlero et al., 2008). The reaction proceed sluggishly (~60% after 1 month) and it is driven by the precipitation of the product. Di-, tri- and tetra-peptides have been formed by this reaction. Despite the low yield, this reaction demonstrate that very simple catalysts may carry out the important function of synthesizing other catalysts (peptides) or—in principle—other molecules starting from simpler building blocks. From the chemical viewpoint, this is not a pure condensation reaction, because there is a leaving group on the carboxylic group. Consider, however, that even today the ribosome catalyses the formation of peptide bonds by nucleophilic attack to the ester moiety of aminoacyl-tRNA.

The *synthetic biology* approach can also provide useful and unexpected insights into the formation of primitive cells. Interestingly, synthetic biology has included as one of its main goal the total synthesis of a living cell, and consequently the research on DNA/RNA/proteins minimal cells has became one of the pillar of this field (De Lorenzo and Danchin 2008). Clearly, such kind of minimal cells has relevance in biotechnology as well as in basic science. From the point of view of origins of life the efforts for the construction of *semi-synthetic* minimal cells (Luisi et al. 2006) will eventually demonstrate: (*i*) that we can learn how a cell works by constructing it, not only by analyzing its components; (*ii*) that a living cell can exists with a limited number of components, at the expenses of efficiency, but yet being recognized as alive; (*iii*) what are the emergent properties of a complex compartmentalized chemical system, i.e., spontaneously arising from the cell organization; and (*iv*)—if still needed—that a living entity can be generated by proper organization of non-living molecules. Clearly, the latter is an example of the emergence of properties that are not present at a lower hierarchical level.

In addition to these long range goals, the efforts to construct minimal cells may lead to intermediate results that may help to understand the origin of protocells. This is the case, for example, of the ever accumulating evidences that compartmentalized reactions differ from bulk ones, being somehow more efficient (Nomura et al. 2003; Cisse et al. 2007). Recently reported evidences suggest that the mechanism of solute entrapment inside vesicles may differ from what expected theoretically (Souza et al. 2009), favouring high entrapment yields in a fraction of the cells population. In particular, attempts to express functional protein in 200 nm (diameter) liposomes brought to the intriguing conclusion that solutes can be entrapped as if they were about 20–50 times more concentrated than the nominal concentration value. Now these preliminary results have inspired a research line on the spontaneous generation of concentrated compartments, clearly related to the problems of the protocell origins.

Of great interest with respect to the structure and function of early biopolymers in primitive cells is the *chemical synthetic biology* approach to the study of short random peptides and nucleic acids (called "never born" because their ortogonality to known sequences), by means of huge library screening (hundred billions different sequences can be screened in search of binding to transition state analogues (Yamauchi et al. 2002), or to finding folded ones (Chiarabelli et al. 2006)). These studies are of paramount importance for defining the plausible structure of early cells, since it is reasonable to think that short weakly-functioning biopolymers (whose synthesis become more accessible in prebiotic terms) preceeded long ones.

In conclusion, although we do not know the structure of simple, early cells, experimental approaches, being synthetic or semi-synthetic, when guided by autopoiesis, can actually give insights and support the origins-of-life oriented studies.

Acknowledgements

This work has been funded by the SYNTHCELLS project (Approaches to the Bioengineering of Synthetic Minimal Cells, EU Grant #FP6043359), by the Human Frontiers Science Program (RGP0033/2007-C), and by the Italian Space Agency (Grant Nr. I/015/07/0). It is also developed within the COST Systems Chemistry CM0703 Action.

Box 1 Nomenclature of minimal cells

Primitive cells, protocells. Simple ancient cells whose structure is compatible with a plausible pathway from prebiotic chemical compounds. Protocells should obey to the continuity principle and therefore should be coherently evolvable to natural existing cells.

Minimal (living) cells. A general term indicating a living cell containing the minimal number of components and performing minimal number of functions in order to be defined as "alive". The "minimal cell" term does not specify the chemical nature of its components. Their molecular structure can be simple or complicated, ancient or modern, natural of synthetic.

Synthetic cells, artificial cells (1). A specific term referring to man-made cells that are not necessarily related to natural cells. They can be built by using whatever compound may perform some specific function. This may lead to ortogonal life, in the sense that these cells have no chemical similarity with natural ones.

Semi-synthetic minimal cells. A specific family of man-made cells constructed by using the minimal number of extant enzymes, DNA, ribosomes, natural or synthetic lipids, and a synthetic (constructive) way of assembly. Although these cells are simpler than natural cells, simplicity of semi-synthetic minimal cells does not imply primitiveness.

Artificial cell (2). A term that is largely used for *in silico* approaches, where artificial cells are structures existing as a computer representation of autonomous, self-reproducing cells.

Autopoietic cell. A general term indicating a cell (of whatever chemical composition) whose dynamics obey to the circular logic of autopoiesis. The autopoietic cell regenerates from the inside all components (boundary molecules included) that are being transformed and/or disposed of, thanks to a network of processes that produces all components, that in turn generate the processes that produce such components, and so on…

Ribocell (RNA-based cell). A very specific denomination indicating a minimal cell based only on ribozymes, that control all metabolic transformations and replicate themselves.

References

Chiarabelli C, Vrijbloed JW, De Lucrezia D, Thomas RM, Stano P, Polticelli F, Ottone T, Papa E, Luisi PL (2006) Investigation of de novo Totally Random Biosequences. Part II. On the Folding Frequency in a Totally Random Library of de novo Proteins Obtained by Phage Display. Chem. & Biodiv. 3:840–859

Cisse I, Okumus B, Joo C, Ha T (2007) Fueling protein–DNA interactions inside porous nanocontainers. PNAS 104:12646–12650
De Lorenzo V, Danchin A (2008) Synthetic biology: discovering new worlds and new words. EMBO Reports 9:822–827
Gil R, Silva FJ, Peretó J, Moya A (2004) Determination of the core of a minimal bacteria gene set. Microbiol Mol Biol Rev 68:518–537
Gorlero M, Wieczorek R, Adamala K, Giorgi A, Schininà ME, Stano P, Luisi PL (2008) Ser-His catalyses the formation of peptides and PNAs. FEBS Letters 583:153–156
Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 93:1–13
Nature Insight (2008) Small-molecule catalysis. Nature 455:303–349
Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. ChemBioChem 4:1172–1175
Rasmussen S, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF, Bedau MA (2004) Transitions from nonliving to living matter. Science 303:963–965
Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390
Varela F, Maturana HR, Uribe RB (1974) Autopoiesis: the organization of living system, its characterization and a model. Biosystems 5:187–196
Yamauchi A, Nakashima T, Tokuriki N, Hosokawa M, Nogami H, Arioka S et al (2002) Evolvability of random polypeptides through functional selection within a small library. Prot Eng 15:619–626

E-mail: stano@uniroma3.it

Synthetic Approach to the Understanding of Minimal Cell Membrane

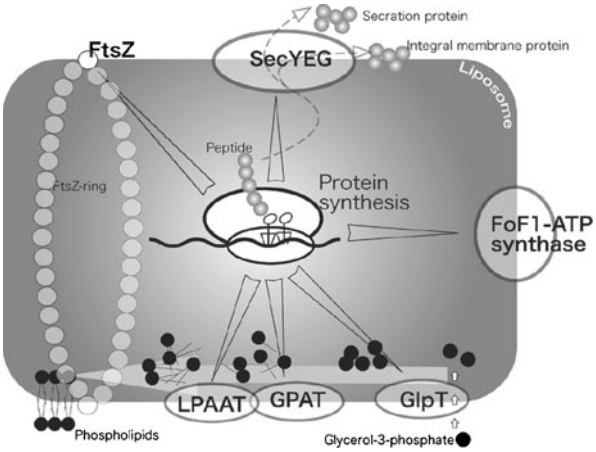
Yutetsu Kuruma

Department of Medical Genome Sciences, University of Tokyo, Japan

Keywords Synthetic Cell • Vesicle Membrane • Cell-free Protein Synthesis • Membrane Protein

Recent progress in biotechnology has accelerated the attempts in constructing a lab-made cell, generally called “artificial cell” or “synthetic cell”. Such research trend has developed as synthetic biology and provided an opportunity to think of the origin of cell (or life) in practical manners. This is based on the idea that the form of a primitive cell might be simple enough, as well as that a prototype cell could be created in the simplest form (Luisi et al. 2006). Thus the construction of a synthetic cell could provide a direct clue for the understanding of the origin of cell. Many theoretical models and basic researches guide the ways in which to achieve the synthetic cell. We are at a stage where the synthetic cell research shifts to the implementation step in biochemical laboratory. Actually, we are able to reconstruct some cellular biochemical systems such as protein synthesis reaction in a test tube. Some of these reactions can even be performed inside a compartmentalized space by means of vesicle manipulation. For instance, Yomo and colleagues have succeeded in synthesizing β-subunit protein of Qβ RNA replicase

within liposome compartments which determines that the synthesize replicase can copy its own template mRNA (Kita et al. 2008).



The next approach to the synthetic cell is the architecture of outer shell of the compartment, i.e., lipid membrane of the vesicle. While it is very important to look into the nature of the lipids constituting the vesicles, it is more so to assemble membrane proteins onto the lipid bilayer in order to give cell-like abilities to the vesicle, e.g. ion-channels and diverse transporters. Regarding this point, we are attempting to synthesize several types of membrane protein using a reconstructed cell-free translation system (PURE system). The PURE system, which consists of 37 purified factors and small molecule components, allows the production of any protein through transcription and translation reactions. Furthermore the PURE system can be encapsulated in lipid-base or fatty acid-base vesicles widely ranging in size: from 100 nm to several hundreds micrometer diameter. In fact, we can easily synthesize protein inside liposome by use of the PURE system and synthetic phospholipids (Kuruma et al. 2009). However, the internal environment of the vesicle has a certain degree of limitation. This is not only due to the limitation of space, but also because the vesicle is a completely closed system which does not allow even a single proton to pass. The present problems can be cited as follows. Firstly, the internal protein synthesis can no longer continue once the internal energy or substrates are exhausted. Secondly, although enzyme is synthesized inside, it cannot work without the supply of its substrate and cofactor from the outside environment. Thirdly, some internally-synthesized products must exit the vesicle (e.g., secretion protein). All these problems are caused by the fact that the vesicle is a tightly closed system. Although the range of permeability of the vesicle depends on the lipid composition of the membrane, several transporter proteins in the lipid bilayer strictly regulate the material exchange across cell membrane.

To investigate this point, we are constructing various kinds of membrane proteins through the internal protein synthesis in vesicle, which seems to play a significant function for the living cell. One example is a set of membrane enzymes: glycerol-3-phosphate acyltransferase (GPAT) and lysophosphatidic acid acyltransferase (LPAAT) involved in phospholipid biosynthesis pathway (3). The GPAT and LPAAT catalyze stepwise acyl-chain binding reactions on the backbone of glycerol-3-phosphate in order to produce a phospholipid. As a result, the newly synthesized phospholipids are accumulated into the lipid bilayer of the mother vesicle. These membrane enzymes must produce sufficient number of phospholipids in order to achieve a self-division of the cell, while it is also essential that the substrate (glycerol-3-phosphate) is continuously

supplied from the surrounding environment. This can be conducted by a specific membrane transporter (GlpT), which is responsible for glycerol-3-phosphate. Another example is FtsZ protein that makes a ring formation along the internal membrane surface of vesicle. Although FtsZ is known as a soluble protein and requires the help of MinCDE system to be attached to the membrane's internal surface, use of a chimeric protein conjugating FtsZ and a membrane-anchoring fragment of MinD protein allows us the direct localization of FtsZ protein onto the membrane (Osawa et al. 2008). This is also a key membrane protein for the construction of a synthetic cell.

Cell membrane has another very important role, namely the energy generation by FoF1-ATP synthase (FoF1). FoF1 is a macromolecular machinery consisting of 8 kinds of protein and synthesizing ATP molecule from ADP and phosphoric acid. The ATP synthesis reaction is catalyzed by F1 complex that is driven by a torque of Fo complex integrated in membrane. Although there are some other energy generating systems (e.g., glycolytic cycle) besides FoF1 in cell, the importance of FoF1 has been mentioned in a theoretical analysis which explores the primitive membrane conditions in early cells (Mulkidjanian et al. 2009). The emergence of an energetically autotrophic cell might be an exact start line of the evolution of modern life.

All the membrane proteins shown in this chapter have been experimentally constructed by our research team. During these attempts, a full set of biological systems for the living cell is (re)constructed by producing each responsible component. The ultimate goal of these challenges is to create a synthetic cell in a living state that fulfills self-maintenance, self-reproduction, and evolution. For this purpose, it is a valid choice to start from the basis of cell-free protein synthesis system since proteins are a major component in the extant cells. Furthermore, determining a minimum set of proteins and genes responsible in sustaining life in a cell is invaluable. This is called Minimal Genome and a prospective cell possessing the Minimal Genome is called Minimal Cell (1). The creation of the Minimal Cell will be the first objective in the study of the synthetic cell and the origin of life.

References

Kita H, Matsuura T, Sunami T, Hosoda K, Ichihashi N, Tsukada K, Urabe I, Yomo T (2008) Chembiochem 9:2403–2410
Kuruma Y, Stano P, Ueda T, Luisi PL (2009) BBA-Biomem 1788:567–574
Luisi PL, Ferri F, Stano P (2006) Naturwissenschaften 93:1–13
Mulkidjanian AY, Galperin MY, Koonin EV (2009) Trends Biochem Sci 34:206–215
Osawa M, Anderson DE, Erickson HP (2008) Science 320:792–794

E-mail: kuruma@molbio.t.u-tokyo.ac.jp

Could the Ribocell be a Feasible Proto-cell Model?

Fabio Mavelli[*1], Pierluigi Della Gatta[1], Luigi Cassidei[1], Pier Luigi Luisi[2]

[1]Dept. of Chemistry, University of Bari, Italy; [2]Department of Biology, University of Roma Tre, Rome, Italy

Keywords Ribocell • Self-replicating RNA • Self-reproducing Vesicles • Artificial Cell

The Ribocell (RNA based cell) is a theoretical cellular model proposed some years ago (Szostak et al. 2001) as a minimal cell prototype. It consists in a self-replicating minimum genome coupled with the self-reproduction of the lipid vesicular container. Szostak and colleagues (2001) envisaged the existence of two hypothetical ribozymes, one (R_{Lip}) able to catalyze the conversion of molecular precursors (P) into membrane lipids (L) and the other (R_{Pol}) able to duplicate RNA strands. Therefore, in an environment rich of both lipid precursors (P) and activated nucleotides (NTP), the Ribocell can self-reproduce if both processes, the genome self-replication and the membrane reproduction (growth and division) would be somehow synchronized. In fact, self-maintenance, self-reproduction and evolvability are traits necessary for a minimal cellular life (Luisi 1998) and Ribocells can exhibit all of them, at least in principle. Therefore, even if this model is highly hypothetical at present, it tries to overcome the conundrum of the plausible simplicity of emergent living organisms, compared to the metabolic complexity of modern cells. In this sense, the Ribocell is much more a genuine prototype of the simplest artificial cell rather than of a primordial cell, since the prebiotic plausibility of the ribozyme synthesis is still under debate. However, when experimentally implemented, it could represent a possible pathway for the transition from non-living molecules to life, showing that abiogenesis is a well-founded hypothesis. On the other hand, from the point of view of the origins of life on Earth, the Ribocell represents the overlap of two different approaches: the RNA-world and the Compartimentalist approach. The former considers the appearance of RNA as a key step in the transition from non-living to living matter, since these biomolecules can exhibit both catalytic (like proteins) and transferring information (like DNA) behaviours; the latter stresses the role of self-assembled amphiphilic aggregates as micro-nano sized compartments where suitable physical and chemical conditions can take place for the emergence of self-maintained and/or self-replicating metabolic networks. In particular, the metabolic products concentration in a vesicle water pool can increase much more rapidly than in the bulk of aqueous solution and this, for instance, can strongly enhance the self-production of biopolymers, which may exhibit lower membrane permeability than the free monomers.

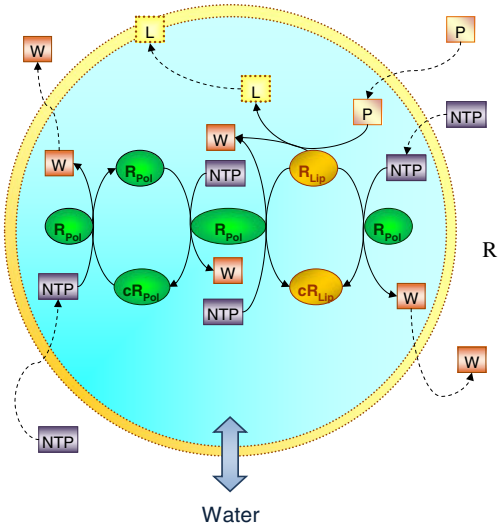
In this contribution we present and discuss a detailed and as realistic as possible kinetic mechanism for the Ribocell. Our aim is to study theoretically the feasibility of the Ribocell as a protocell model and, at the same time, give insights and clues to researchers who are involved in its chemical implementation in a test tube. All the needed substrates are assumed to be available in the external environment and the two ribozymes are already present and encapsulated in the lipid vesicle, i.e. starting from a first ancestor. In fact, the encapsulation of large biomolecules into lipid vesicles has been described and experimentally achieved by dehydration-hydration cycles (Shew and Deamer 1983), while a new synthetic strategy (Powner et al. 2009) makes the RNA production in a prebiotic environment more plausible. Therefore, the main question we try to answer is if synchronization can emerge spontaneously by coupling metabolic reactions, i.e. the lipid production and the genome duplication, along with chemical-physical processes (e.g.: solute transport across the membrane, water flux driven by osmotic pressure gradient or elongation and bending of the lipid bilayer). In other words: is the first ancestor able to reach a self-maintaining/self-reproducing regime oscillating continuously between two stationary states before and after the division?

The Ribocell kinetic model

The kinetic model proposed to describe the dynamic behaviour of the Ribocell (Fig.1) is consistent with a model recently used for reacting vesicles (Mavelli & Ruiz-Mirazo 2007a) and applied to the simulation of the competitive processes taking place in

aqueous solutions of oleic acid and POPC vesicles (Chen et al. 2004; Cheng and Luisi 2003; Mavelli & Ruiz-Mirazo 2007b; Mavelli et al. 2008). A reacting vesicle is treated as a homogeneous reacting aqueous domain with a time variable volume V_C that can exchange through the membrane water and substrates with the external environment. These molecular fluxes are driven by osmotic pressure and concentration gradients according to the bilayer permeability. Moreover, the membrane can associate and release lipid molecules from and to the aqueous solution and its energetic state can be monitored by means of the reduced area: $\Phi = S_\mu / \left(36\pi V_C^2\right)^{1/3}$ (S_μ being the actual surface area). For a perfectly spherical vesicle Φ equals 1.0, while $\Phi<1$ if the lipid aggregate is swollen and the membrane is in an elastic tension state that can lead to burst (osmotic crisis). If $\Phi>1$, the vesicle is deflated and we assume it spontaneously divides when two twin spherical aggregates can be formed, i.e. if $\Phi=\sqrt[3]{2}$ (Mavelli & Ruiz-Mirazo 2007a; Mavelli et al. 2008).

On the right of Figure 1, the internal metabolism adopted for the Ribocell is reported in details. Both pairs of RNA strands reversibly associate (1) and these processes are shifted towards the dimer formation and strongly dependent on temperature.



$$R_{Pol} + {}_cR_{Pol} \underset{k_S}{\overset{k_{SS}}{\rightleftharpoons}} R_cR_{Pol}$$
$$R_{Lip} + {}_cR_{Lip} \underset{k_S}{\overset{k_{SS}}{\rightleftharpoons}} R_cR_{Lip} \quad (1)$$

$$R_{Pol} + S \xrightarrow{k_{R@S}} R@S$$
$$R@S_{cn}S + NTP_{n+1} \xrightarrow{k_{NPT}} R@S_cS_{n+1} + W \quad (2)$$
$$R@S_cS \xrightarrow{k_{R@SS}} R_{Pol} + SS_c$$

$$P + R_{Lip} \xrightarrow{k_L} L + R_{L\,ip} + W \quad (3)$$

Fig. 1—The Ribocell kinetic model: graphical representation of membrane transport processes (dashed lines) and internal reactions (solid lines) on the left, internal metabolism on the right. (1) reversible association of strands of RNA polymerase (R_{Pol}) and RNA-synthase (R_{Lip}) with their complementary strands ${}_cR_{Pol}$ and ${}_cR_{Lip}$, (2) catalytic cycle of the replication of RNA strands (S=R_{Pol}, ${}_cR_{Pol}$, R_{Lip} and ${}_cR_{Lip}$), (3) conversion of the precursor P into the membrane lipid L catalyzed by the ribozyme R_{Lip}

The replication of any RNA strand is catalyzed by the polymerase R_{Pol} according to pair base linking mechanism reported in bracket (2). The process starts with R_{Pol} binding any monomeric strand S to form the complex R@S. Then this complex will initiate the polymerization of the conjugate strand ${}_cS$, by coupling iteratively the complementary bases and releasing the byproduct W. When the strand ${}_cS$ has been completely formed, the polymerase releases the new dimer. Finally, the precursor P is converted into the lipid L by the assistance of the ribozyme R_{Lip} (3).

Table 1 Kinetic Constants and Permeability of the Ribocell at room temperature (S=R_{Pol}, $_cR_{Pol}$, R_{Lip} and $_cR_{Lip}$).

Kinetic Constants	**Values**	**Description**	**Reference**
$k_{SS}[s^{-1}M^{-1}]$	$8.8{\cdot}10^{6}$	Formation of R_cR_{Pol} and R_cR_{Lip}	[Christensen 2007]
$k_S[s^{-1}]$	$2.2{\cdot}10^{-6}$	Dissociation of R_cR_{Pol} and R_cR_{Lip}	[Christensen 2007]
$k_{R@S}[s^{-1}M^{-1}]$	$5.32{\cdot}10^{5}\times10$	Formation of R@S	[Tsoi & Yang, 2002]
$k_{R@SS}[s^{-1}]$	$9.9{\cdot}10^{-3}$	Dissociation of Complexes $R@S_cS$	[Tsoi & Yang, 2002]
$k_{NTP}[s^{-1}M^{-1}]$	0,113	Nucleotide Polimerization in Oleic Vesicle	[Mansy et al., 2008]
k_L $[s^{-1}M^{-1}]$	$0{,}017\times10^{7}$	Reaction catalyzed by the Hammerhead ribozyme	[Stage-Zimmermann & Uhlenbeck, 1998]
k_{in} $[dm^2s^{-1}]$	$7.6{\cdot}10^{19}$	Oleic acid association to the membrane	[Mavelli et al.2008]
k_{out} $[dm^2s^{-1}]$	$7.6{\cdot}10^{-2}$	Oleic acid release from the membrane	[Mavelli et al.2008]
Permeability $[cm{\cdot}s^{-1}]\times10^{-8}$	**Values**	**Description**	**Reference**
P_P	0.42	Membrane Permeability to Lipid Precursor	
$P_A=P_U=P_C=P_G$	0.0019	Oleic Acid Membrane Permeability to Nucleotides	[Mansy et al., 2008]
$P_W=P_T=P_R$	0.0	Membrane Permeability to W and Ribozymes	
P_{water}	$1.0{\cdot}10^{6}$	Oleic Acid Membrane Permeability to Water	[Sacerdote & Szostak, 2005]

In order to duplicate the entire "genome", the replication cycle should operate on all the four RNA strands at least once. This requires that at least 3 strands in a vesicle must be present simultaneously: R_{Pol} necessarily, a filament of R_{Lip} or $_cR_{Lip}$ and another R_{Pol} or $_cR_{Pol}$. The genome duplication has also the effect of accumulating inside the vesicle the waste W, usually a charged leaving group with low membrane permeability. This can force a flux of water from the outside that can bring the vesicle in an elastic tension state $\Phi<1$ if it is not properly counterbalanced by a membrane surface increase due to the catalyzed production of lipids.

Deterministic Results

The time behaviour of the Ribocell can be obtained from the previously described kinetic model by using a stochastic or a deterministic approach (Mavelli & Piotto 2006). In this contribution deterministic calculations are used to follow the average time course of a vesicle population by solving the ordinary differential equation set associated with the ribocell metabolism (for further mathematical details see: Mavelli, forthcoming). The two ribozymes are assumed both 20 nucleotides long (the minimum number for observing a folded RNA conformation) with a random sequence of bases and, for the sake of simplicity, with a similar kinetic behaviour, so that the exact sequence of the RNA strands can be neglected.

Table 1 shows the values of the kinetic constants and membrane permeability used with the respective references from which they were derived. All values are considered at room temperature. It is worthwhile to remark that the constant $k_{R@S}$ for the R@S complex formation and the constant k_L for the catalyzed synthesis of lipids have been enhanced of factor 10 and 10^7 respectively. The only parameters assigned arbitrarily are then the membrane permeability to the byproduct: P_W=0.0 cm/s, based on the assumption that W is a charged species, while the

permeability to the precursor: P_P=0.42·10^{-8}cm/s is comparable to those of several organic compounds (e.g.: oleic acid membrane permeability to Arabitol—Sacerdote & Szostak 2005).

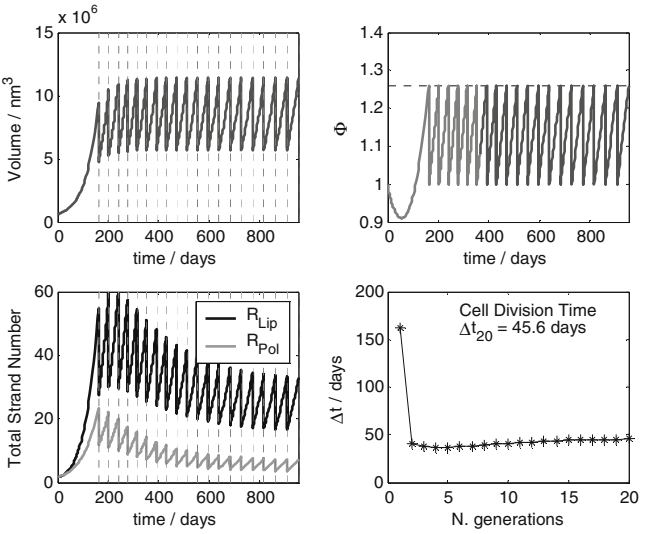


Fig. 2—Ribocell deterministic time evolution: on the top row volume (left plot) and the reduced surface (right plot); on the bottom row the population of ribozymes against time (left plot) and division times against number of generations (right plot). Vertical lines mark the division times while the horizontal line represents the dividing condition ($\Phi=2^{1/3}$). The simulation was made for an initially spherical ($\Phi=1$) oleic acid vesicle with a 50 nm radius. At time zero only two ribozyme dimers R_cR_{Pol} and R_cR_{Lip} are present in the vesicle core, internal and external concentrations of all nucleotides and precursor P were put equal to 500 μm; [W] was set 0.0 and [L] equal to the equilibrium value 66.7 μM [Chen et al.,2004]. At time zero, the vesicle was in a perfect osmotic balance

Figure 2 shows the time evolution of the Ribocell in a time range of 2.6 year showing how a cyclic regime of growth and division takes place. After 10 generations, the system reaches a stationary condition where the Ribocell doubles its molecular content and its size during the growth phase, and then it generates two twin daughter cells with spherical radius of about 110 nm. The graph in the bottom right corner shows the trend of time interval Δt between two successive divisions that stabilizes around 45.6 days after about 10 generations. The average populations of ribozymes are reported in the lower left plot of Fig. 2. At the end of any growth cycle the total genetic material doubles and when the stationary condition is reached after each division the R_{Pol} and R_{Lip} total numbers are constantly equal to 16 and 6 units respectively, showing behaviour perfectly synchronized with that of the membrane. The observed difference between the two populations can be ascribed to the lower availability of R_{Pol} strands to act as templates being involved as catalysts.

Conclusions

In this paper a kinetic model for the Ribocell was presented in order to study its deterministic time behaviour by using kinetic constants and permeability values assigned from data in literature. Although this work is still in progress, nevertheless it shows that synchronization between genome duplication and membrane reproduction is possible, within the approxima-

tions and the kinetic parameters used. The high value of reproduction time obtained: 45.6 days, compared to that of real cells, can be ascribed mainly to the low value used for the dissociation constant k_S of RNA dimers at room temperature. In fact, an increase of the work temperature could greatly increase the efficiency of reproduction as shown recently by the process of self-catalyzed replication of RNA strands (Lincoln and Joyce 2009). Finally, it is important to remark that the catalytic efficiency of the R_{pol} ribozyme appear to be a crucial parameter since the k_L value used for the lipid production is comparable to that of modern enzymes.

References

Chen IA, Roberts RW, Szostak JW (2004) Science 305:1474–1476
Cheng Z, Luisi PL (2003) J Phys Chem B 107:10940–10945
Lincoln TA, Joyce GF (2009) Science 323:1229–1232
Luisi PL (1998) Orig Life Evol Biosph 28:613–622
Mansy S, Schrum JP, Krishnamurthy M, Tobe S, Treco DA, Szostak JW (2008) Nature 454:122–126
Mavelli F (forthcoming) In: Luisi PL (ed) The minimal cell, Springer
Mavelli F, Piotto SJ (2006) Mol Struct 771:55–64
Mavelli F, Ruiz-Mirazo K (2007a) PhilTrans R Soc B 362:1789–1802
Mavelli F, Ruiz-Mirazo K (2007b) In: Bezrukov SM (ed) Noise and fluctuations in biological, biophysical, and biomedical systems. SPIE (United States), Bellingham, 6602, pp 1B1–1B10
Mavelli F, Lerario M, Ruiz-Mirazo K (2008) In: Arabnia HR et al (eds) BIOCOMP'08 Proceedings (Vol II), CSREA, 934–941
Powner MW, Gerland B, Sutherland JD (2009) Nature 459:239–242
Sacerdote MG, Szostak JW (2005) PNAS 102:6004–6008
Shew R, Deamer D (1983) Biochim Biophys Acta 816:1–8
Stage-Zimmermann TK, Uhlenbeck OC (1998) RNA 4:875–889
Szostak JW, Bartel DP, Luisi PL (2001) Nature 409:387–390
Tsoi PY, Yang M (2002) Biochem J 361:317–325

E-mail: mavelli@chimica.uniba.it

On the Minimal Cell

Pier Luigi Luisi

Department of Biology, University of Roma Tre, Rome, Italy

Keywords Protocells • Synthetic Cells • Cell Models

The project "minimal cell", as initiated in the early 90ties at the Swiss Federal Institute of Zurich (ETHZ) by the work of Schmidli et al. (1991) and Oberholzer et al. (1995), had two complementary aims: clarify what is the minimal and necessary complexity to have a minimal living biological cell (let us call this MLC); and to clarify the structure of the early cells at the beginning of life, let us call them simply protocells. The two starting points for this inquiry are the questions of whether the high complexity of modern cells (often thousand of genes) is really necessary for cellular life, and the consideration that early cellular life could not have conceivably started right away with such a degree of complexity.

It took 10 years or so before this idea of MLC became accepted as a viable experimental project, generally using vesicles as model cellular compartments. Nowadays there are several groups dealing with the question of the minimal cell.

What has been achieved thus far? And what are the present stumbling blocks?

The most important result, I believe, is the achievement of protein synthesis inside vesicles, often the green fluorescence protein (GFP) for obvious detection reasons. This has been achieved with systems based on one hundred or so genetic components, thus a considerable reduction of complexity with respect to modern cells.

On the other hands, two major features of living cells have not been reconstructed yet. In fact, these systems usually work as one pot reactors, namely they produce GFP in one batch reaction and then stop. Thus, self-sustainability, namely the maintenance of a stationary continuous system, has not been reached yet. The problem here is given by the very poor permeability of the vesicles, so that the metabolites possibly added externally do not flow inside the vesicles.

The other feature which is missing is the self-reproducibility, namely the capability of these semi-synthetic cells to re-make themselves, which implies a core-and-shell replication. Those are two serious limits, also because is not at all apparent how they can be overwhelmed.

For the first problem, one may think of putting porine or some other channel proteins on the membrane, and this has been tried. But this is not a prebiotic solution. More prebiotic would be the use of a very permeable vesicular system-but this would also be very leaky, with corresponding problems of stability of the whole system. Concerning the problem of the cellular self-reproduction, one would need to modify the encapsulated ribosomal system—so that it makes copies of all the enzymes—but this would of course increase the complexity of the system. Then we are back to a rather large number of genes…

How, is then possible to pursue research on MLC and on the early protocells at the same time, if one has to recognize the necessity of 100 genes or so for minimal cellular life-a degree of complexity which is not consistent with the very early stage of prebiotic development-the age of the protocells?

There is an important conclusion which arises from this question. This is the necessity of disentanglement of the notion of MLC from the notion of proto-cells. In other words, if the minimal living cell, with the trilogy of the three properties, is going to encompass, say, one hundred genetic products, then it cannot represent at the same time the structure of the early protocells. We have therefore to conclude that the two projects cannot be unified. The very early protocells and the MLC should then be seen as two distinct concepts, two distinct projects, two distinct structures.

Whereas it is right to pursue the search for the minimal living cell using 100 or so genetic components, we have to re-start thinking about protocells, and how can they be made partially functional. We have to conceive protocells which are not living and that proceed towards MLC only through a complex pathway of molecular evolution.

For this new protocell project we should forget self-reproduction, which is a much later development. Instead, one should focus on the synthesis of very simple polypeptides (and not proteins with 20 different amino acids); and focus on some corresponding simple form of metabolism. At this level, one should for example think in terms of primitive forms of ribosomes. One example of such attempts is offered by the work carried out by Chris Thomas, a former PhD student of my group, who studied the interaction of rRNA with poly-L-arginine. The starting consideration is that basic peptides may play a similar structural role of ribosomal proteins, due to their favourable interaction with anionic ribonucleic acid. Preliminary experiments have shown that rRNA/poly-L-arginine complexes form rapidly and spontaneously by simple mixing the two components, in definite molar ratios. Surprisingly, the resulting complexes show a compact structure as evident by cryo-TEM imaging and dynamic light scattering, and have similar dimension and gross form of ribosomes. Presently, these studies are pursued in our

laboratory in order to see whether these simple complexes may be provided of some activity. If successful, this investigation will reveal that it is indeed possible to hypothize primitive ribosomes composed by RNAs and simple peptides kept together by simple physical forces like electrostatic and hydrophobic ones, as well as hydrogen bonds and therefore suggest a possible origin of ribosomes as a primitive complex of complementary molecules. In more general terms, it seems to me that the research project on protocells must be re-thought and re-designed from the very beginning, and that it should be no longer confused with the project MLC.

References

Oberholzer T, Wick R, Luisi PL, Biebricher CK (1995) Enzymatic RNA replication in self-reproducing vesicles: an approach to a minimal cell. Biochem Biophys Res Comm 207:250–257

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

E-mail: luisi@mat.ethz.ch

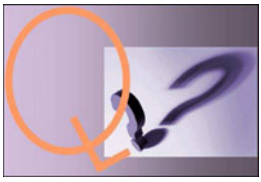
Workshop OQOL'09

Extended Abstracts for the Following Selected Question

- **Life as unity or confederacy?**

Premise. Many sciences have conventionally (if implicitly) referred to "life" as a unitary concept, and all too often, we speak of "the origin of life" as if it were essentially one kind of unified event: a transition from "no life" to "life" on Earth. An alternative premise would be that life is a collection of coupled but still distinguishable subsystems, each with its own recognizable dynamics and requirements for stability. In that case the origin of life could involve a sequence of transitions understandable in somewhat independent terms. For instance, one could take separately the appearance of self-reproducing systems and the formation of vesicles, biogenesis of proteins different from setting up metabolic cycles, origin of reductive power different from prebiotic chemistry, etc The degree of both contingency and of what some have called "irreducible complexity" in life will depend strongly on how tightly or loosely its subsystems are coupled.

The question. Do you agree with this possible alternative view of life origin? And if yes, what is the proper way to apply the notions of interdependency versus subsystem independence, in the understanding of both the modern function of life and of its origin? Can a different understanding of the organization and stability of life today lead to better sequences of investigations of life's origins? If "life" is not a totally unitary notion, but rather a confederacy of coupled processes, can the recognition of this decomposition help us define the nature and process of origins of life in ways that do not lead to contradiction and confusion?



Confederacy Based on Synthetic-pathway and Bio-energetic Modularity

Eric Smith, Harold J. Morowitz

Santa Fe Institute, 1399 Hyde Park Road, Sana Fe, NM 87501; Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA 22030

Keywords Autotrophic Origins • Modularity • Bioenergetics • Biosynthesis

Unitary origin of life versus origination through modules

We begin with the observation that all experimental and theoretical approaches to origins of life, which attempt to study the emergence of order in isolated components of the living state, are de facto commitments to describing life in terms of modules. Laboratory conditions or hypothesized early worlds provide contexts sufficiently less complex than modern cells to be tractable to us, in which the requirements for order within different subsystems can be studied independently. If we assume that lessons learned about subsystem organization remain informative about order in living systems, we have assumed that in real life the subsystems interact either sufficiently weakly or through sufficiently simple interfaces that the internal subsystem dynamics are not qualitatively unlike those we can reproduce.

All of these assumptions can be doubted, since we do not observe the heterogeneous, refined, particular chemical structures and processes of life separately in any non-living system, or in any less complex contexts than fully integrated autotrophic bacterial or archaeal cells.[1] Yet as a practical matter the study of modules is unavoidable. Moreover, the assumption that life has a modular architecture and that it emerged in a sequence of stages has strong likelihood (in a Bayesian sense) based on our understanding of random processes, and it is the only one to which scientific method can be applied. The likelihood argument has been cleanly presented by Simon (1973): the only route for stochastic processes to produce structures of considerable complexity from unstable elementary steps (such as those deviating from thermodynamic equilibrium), without invoking events of extreme improbability, is through formation of modules which are stable intermediate points of construction.

The two central questions in choosing experiments or theories are then: 1) which subsystem boundaries will be most informative about the major transitions that led to life as we now see it, and 2) does a large enough quantitative difference exist between intra- and inter-system interactions to permit useful study of modules in simplified contexts? The likelihood (in the sense of Simon's argument) for a putative modular description depends not only on tractability of study, but on the probability that the proposed modules could have been stable holding points for a stochastic process in conditions plausible for the early earth.

From this perspective the difference between genes-first and metabolism-first approaches to origins arises from whether individuals and population dynamics, or physiological and ecological universals, are preferred as grounds for decomposition. The genes-first approach draws on a heavy (if often unstated) influence of the Central Dogma in thinking about molecular biology, and of population genetics on thinking about evolutionary dynamics. These two thought systems mesh coherently, because both take the Darwinian individual (capable of preserving heritable

[1] If ecosystem as well as organism contributions to complexity are seriously acknowledged, heterotrophy leads only to more complex contexts. In this sense, autotrophic cells may be understood as wholly contained ecosystems.

variation, replicating, and perishing) as the natural unit, with information input through selective processes at the population level, and phenotypes produced more or less mechanistically from genotypes. Limitations of this point of view are that it gives a somewhat impoverished treatment of the constraints of development and ecology which may structure the space of possible phenotypes a priori, and its need to assume the existence of a quite complex entity capable of functioning as a Darwinian individual. To some degree the latter complexity can be displaced from the individual to its context, as when the replication-selection model is applied to RNA oligomers in an Eigen hypercycle, competing for activated monomers in a primordial soup.

The contrary feature of metabolism first—as we will represent the idea here—is a diminished emphasis on both the individual and any distinctively Darwinian dynamic, and much stronger emphasis on coherence in biosynthetic pathways across organisms, spanning the organism/ecosystem distinction, and connecting biochemistry to geochemical context. As for genes-first paradigms, we will argue that feedbacks between modules essential to maintaining modern cellular life could have been weak or absent in earlier life. However, rather than draw a boundary between a controller (RNA) and a controlled substrate (the metabolome) and remove the feedback by supplying the metabolome exogenously (a primordial soup), we will argue that the feedbacks between universal biosynthetic modules arise when these modules form *micro-environments* for each other energetically and chemically, and we will propose that it was the support from such environments which was originally provided exogenously and perhaps independently by geochemical environments.

A further test for the appropriateness of a modular decomposition of the living state is that modules proposed to be primordial should retain some recognizable autonomy from one another in extant life. The autonomy may be manifest as partial independence from control through hierarchical systems (such as the genome or descent within lineages), or as independent response to environmental pressures that is regular across systems. Partial autonomy is to be expected as more than a vestige of a past state. A corollary to the Simon likelihood argument is that modules are effective routes to complexity precisely because *they do not depend* on integration within higher-order systems for their stability. Instead, they contribute autonomous stability to the systems comprising them, reducing the complexity of the higher-order assembly problem (while at the same time limiting the forms it can take).

Both genes-first and metabolism-first decompositions meet this criterion, although in different ways. Support for a division between the genome and cell physiology can be drawn from the partial independence of viral or transposon dynamics from host-cell lineages. For our metabolically rooted modularity, the support is physiological and evolutionary plasticity between pathway networks and cellular energy systems: for example, the way diverse and heterogeneous redox couples are converted to a universal currency of phosphate esters through the machinery of oxidative phosphorylation, or the way sugars have been substituted for amino-sugars as structural elements in nitrogen-limited plants versus nitrogen-sufficient bacteria. An important question about extant life which would contribute to assessing the relative likelihoods of gene-first versus metabolism-first decompositions—which we do not yet have tools to answer—is how much of the stability of biosynthetic pathways owes to feedback through population-genetic mechanisms depending on the hierarchy of genomic control, and how much reflects environmental constraints that preclude other solutions, perhaps expressed through a form of absolute normalizing selection or lack of evolutionary divergence.

We believe that explicitly regarding life as a confederacy, and integrating the study of its modules with attempts to reproduce stages of origin, reframes existing problems and opens new ones in useful ways. It emphasizes diverse contributions to the stability of the living state, which range from chemical kinetics to trophic ecology. It suggests that the emergence of individuality—of different kinds, at several events—is only one aspect of the emergence

of life, but it is the particular aspect that mediates the emergence of Darwinian dynamics as distinct from other forms of geochemical self-organization. Finally, it suggests that we must explain both the existence and the limits to hierarchies of control systems, and that retracing the difficulties of origins may be a useful way to do so.

Suggestions of modular organization in modern life

Descriptions of life as hierarachical or modular may be made in many ways. We will emphasize modules which follow common divisions in the network topology, chemistry, and energetics of core biosynthesis, and which are then recapitulated at higher levels of cellular organization and physiology, and evolutionary diversification. In all cases we propose modules at the deepest level which appear to be universals of all known life, suggesting either that they antedate all forms of adaptive variation and were "frozen" into the structure of life by the dependence of other systems upon them, or that they are unique solutions to certain problems of function within a self-maintaining system. In the latter case they would define a grey area where strongly convergent evolution becomes indistinguishable from "physical constraints" against evolutionary variation.

The network topology of core biosynthesis suggests a natural decomposition into reaction groups with high internal inter-dependence, separated by "gateway" molecules or reactions. Following the "robust, yet fragile" topology observed for a minimal biosynthetic chart derived from *Aquifex aeolicus* (Srinivasan and Morowitz 2009a, 2009b), we suggest the following as functionally integrated modules. Fig. 1 shows the major reactions in four carbon fixation pathways and a subset of the universal biosynthetic chains. 1) Citric-acid or TCA cycle arcs or loop for arriving at core carbon skeletons. 2) The gluconeogenic pathway connected to the TCA reactions through pyruvate or phospho-enol-pyruvate, with 3) the reductive pentose-phosphate (Calvin-Benson) network as an elaboration of aldol and retro-aldol condensations about the gluconeogenic (or its reverse, the glycolytic) pathway. (Note that whether or not the Calvin-Benson pathway is used for carbon fixation, several of its reactions are key steps in the synthesis of electron-transfer cofactors and aromatic amino acids, and its aldol-reaction steps form the network for producing ribose.) 4) The fatty-acid synthesis pathway from malonate and 5) the isoprenoid synthesis pathway from acetoacetate represent two recursive chain-elongation pathways linked to the TCA reactions through acetate. The biosynthesis of amino acids is more complex but still regularly structured, clustering into simple syntheses from citric-acid intermediates with small number of reactions, and more complex syntheses drawing on more remote modules of the metabolic chart such as the module through chorismate for the aromatics (Copley et al. 2005; Srinivasan and Morowitz 2009a).

Redundant chemistry, chemical energetics, and serving cofactors often align with the network modules. Reduction reactions within the TCA cycle take CO_2 to acetate, an exergonic transformation. Further exergonic reductions then begin from acetate in fatty acid and isoprene synthesis, linking the α-keto-acid and lipid networks energetically (Smith and Morowitz 2004. Thioester to phosphate substrate-level phosphorylation (de Duve 1991) at two key steps (carbonylation of acetyl-CoA and of succinyl-CoA) is a key feature of redox-to-phosphate energy conversion in the TCA reactions. Phosphorylation is a distinctive feature of the gluconeogenic pathway and the aldol condensations that branch from it, often removing OH groups from the network of available aldol condensations, as when glyceraldehyde-3-phosphate and dihydroxyacetone phosphate condense to fructose-1-6-bis-phosphate. A recurrently used small set of phosphoryl transfers, amino transfers, and reductive aminations, used in the early stages of amino acid biosynthesis, is remarkably stereotypically arranged in the same patterns as base assignments in the genetic code (Copley et al. 2005).

Cofactors, which mediate the recurrently used organic reactions as catalysts or group-transfer agents, are correspondingly used unequally across modules. Coenzyme-A functions in thioester formation in some TCA reactions and fatty acid synthesis. ATP as phosphoryl donor is active in dehydrations. NAD is the primary source of reductants, with flavins and deazaflavins secondary. Ammonia enters the organic nitrogen cycle almost exclusively through the formation of glutamate and glutamine, which then act as amine transfer agents in most secondary aminations. S-adenosyl-methionine, folates, and pterins act as carriers of C_1 groups in various states of reduction.

We attach significance to the chemical and bio-energetic divisions among modules because they mimic divisions among energy sources and energy carriers in geochemical processes, particularly hydrothermoal processes acting at the interface of the hydrosphere with the tectonically active lithosphere. Geothermal mantle convection brings reduced metals into contact with seawater, producing copious reductant, and volcanic activity produces dehydrated phosphates at least in surface environments. Strong pH gradients are produced in most hydrothermal systems, with vent effluents ranging from highly acidic for magma-hosted systems to highly basic for peridotite-hosted systems (Martin et al. 2008). Thus biochemistry has preserved the distinctive chemical structure-forming capacities of the three major geological energy sources.

The module boundaries we have suggested may be seen again (more weakly) in signatures of evolutionary conservation. Elaborations of carbon fixation (six forms are now known; four are shown in Fig. 1) have conserved the core modules. TCA arcs run either as a loop autocatalytically in reductive TCA organisms, or in parallel along oxidative and reductive branches in acetogens and methanogens, suggesting that a commitment to these precursors was formed before divergence of these two deepest-rooted carbon fixation strategies. The 3-hydroxypropionate pathway parallels the reductive TCA arc from a precursor in malonate that is the gateway to fatty acid synthesis, and then uses the glyoxylate bypass to reach other synthetic precursors (Lengeler et al. 1999). As noted, the Calvin-Benson network reverses the direction of the gluconeogenic pathway to feed reduced carbon into TCA reactions via pyruvate. Enzyme conservation across clades shows similar boundary distinctions. Attempts to use phylogenetic weighting to arrive at principled likely gene inventories in the LUCA (Mirkin et al. 2003, Fig. 8) show strikingly little variability for enzymes governing core biosynthesis from acetate, and much greater variability for chemically homologous reactions from succinate, recapitulating the apparently very old divergence between TCA and acetyl-CoA carbon fixation, where the loop is either maintained or broken at this point.

Finally the energetic, geometric, and topological organization of the cell recapitulates in many respects the modularity of core biosynthesis. Cellular energy is carried on three systems: redox couples, proton-motive-force, and phosphate esters. Substrate-level phosphorylation has been proposed (de Duve 1991; Martin and Russell 2003; Martin and Russell 2006) as the earliest direct coupling between redox and phosphate energy carriers, but it is a mechanism dependent on compatible bond structure and energies. More flexible coupling is now mediated by protons, which form a “classical” energy currency decoupling the quantum transitions of oxidation/reduction and phosphoryl transfer. This system is entirely dependent on the capacitance and proton insulation/semiconduction chemistry and geometry of cell membranes, and on the topology of cellular compartments (*e.g.* the periplasmic space) to maintain pH and voltage differences.

A proposal for modular origins

In the modular organization that we have suggested governs extant life, the biosynthetic modules appear as both the simplest and the most exclusively self-referential, in that they require from higher levels only the provision of catalysts and the orchestration and

buffering of energy systems. (Even for energy systems, some degree of independence is attained through substrate-level phosphorylation.) Universality at all higher levels depends on that in the biosynthetic core, directly for the cofactors and mediated by more complex constructions for cell form and physiology. We therefore interpret the modularity recapitulated across the hierarchy as one rooted in and constrained by biosynthesis. Because the biosynthetic modularity in turn mimics modularity in geochemical energy sources, we interpret it as the preservation of a transition stage between geochemical organization and the first biochemical organization.

We suggest that core biosynthetic pathways first formed as geosynthetic pathways dependent separately on externally provided chemical energy sources. Because the synthesized organics are not long-lived, these groups of reactions must have sometimes occurred in the same places, but their energy sources need not have originated locally, if these could be preserved and transported in mineral substrates (G. Cody, pers. comm.). The emergence of biochemistry as a distinct form of organization came when these modules formed "micro-environments" for each other, exchanging the dependence on exogenous geochemical sources for dependence on their collocation and coupling. The most explicit such proposal is for the coupling of redox and phosphate energy systems in the earliest protocells, which would have freed biosynthesis from anhydrous phosphate—a short-lived reagent once exposed to seawater—and permitted it to depend only on the more diverse and more ubiquitous environmentally available redox couples, as cells do today (Lengeler et al. 1999).

A biochemical modularity originally coupled to geological processes suggests a path to an RNA world in which RNA catalysts could originally be selected as replacements for prior mineral catalysts (Copley et al. 2006). Some form of template-directed replication would be necessary for the preservation of sequence information, but the primary determinant of fitness could have been increased capacity of pre-existing biosynthetic subsystems, rather than Darwinian competition among RNAs for self-replication within an externally supplied resource pool. The shift in emphasis we propose places greater separation between mechanisms for molecular replication and determinants of fitness (an idea also recently pursued by Nowak and Ohtsuki (2008), and suggests that the most plausible path to an incipient RNA world would have been one in which RNA catalysts preserved the structure of core biochemistry rather than over-wrote it. In this sequence an RNA replication system need never have been autonomous, but could have co-evolved with polypeptide and biosynthetic machinery from the start. While it is analytically more complex to consider, it is probably more realistic to expect that biochemical organization formed through a bootstrapping sequence of replacement of catalysts—and perhaps less frequently of pathways or core reagents—with the major classes of small-molecule constituents such as cofactors and aptamers taking on their key functional roles during this transition toward an RNA world (Copley et al. 2006).

Similarly, cellular organization within this interpretation would have been driven by the energetic and catalytic advantages of compartmentalization, and only later taken on aspects of Darwinian individuality. The view of a late-emerging cellular individuality, and one partly independent of the individuality of RNA sequence lineages, is compatible with the notion of a progenote advocated by Woese (Woese 2000; Woese 2002; Vetsigian et al. 2006), and with the modern understanding of the relation of viral to free-living cell lineages. It also suggests that biofilms, whether self-created or partly exogenous (discussed in our other abstract) may have been important from the very earliest transition from geochemical to cellular organization (J. Baross, pers. comm., see also Woese 1998; Hausner and Wuertz 1999; Schrenk et al. 2004; Chia et al. 2008; Brazelton and Baross 2009), an idea compatible with proposals for mineral-to-cell transitions (Martin and Russell 2003).

However, the most plausible route to the emergence of cells remains very unclear, because phase-separation could have been an early mechanism leading to macromolecular synthesis (Mulkidjanian et al. 2009; Pohorille 2005), and phylogenetic analyses suggest that all major membrane systems were present in the organism or community that constituted the LUCA and subsequently differentiated into the major domains (Pereto et al. 2004).

Tests for residues of primordial modularity

The search for biosynthetic modules to be reproduced in laboratory experiments can be guided in part by reconstructions of plausible earth geochemistry, and in part by testing extant cells for partial subsystem independence. The latter tests may be performed physiologically (experimentally or with flux-balance analyses) or through phylogenetic inference of the history of organism and ecosystem compositions.

Physiological tests for independence of subsystems within the biosynthetic network may look for the minimum level of regulation through gene expression or allosteric enzyme response preserving organism viability, or alternatively for autonomous response of biosynthetic sub-networks to changes in energy carriers and substrate concentrations. Minimally regulated states may be sought, together with reduced but still viable biosynthetic networks, by global transposon mutagenesis (Hutchison et al. 1999). as part of ongoing efforts to produce minimal orgnisms. For enzymes in which allosteric response to substrate levels is part of a regulatory system, a more complex targeted mutation might be introduced to maintain catalytic function at the active site but to reduce response to signals.

At the same time as regulation mechanisms can be removed to see whether cells can remain viable without them, cell phenotype can be tested for response to changes in environmental condition that are independent of active regulation. Growth-rate and flux-model responses to changes in medium composition may be compared. In both studies, numerical knockout in models may be used cooperatively with experimental minimization to see how the degree of independence in subsystem response varies with whole-network complexity. We find autotrophs preferable model systems for such studies, because they are metabolically complete.

Metabolic reconstruction from the metagenomes of ecosystems, which might be termed *meta-metabolomics*, offers a comparative mode of analysis complementary to physiological studies. We may ask to what extent autotrophic organisms are stereotypical as self-contained ecosystems, and whether there are constrained elements within biosynthesis that have been completely preserved over the course of evolutionary history. To understand the degree of plasticity of ecosystem-level metabolism, and so interpret observations of total preservation, we may reconstruct the routes by which pathways have been gained or lost by organisms (Kreimer et al. 2008; Borenstein et al. 2008), and understand the constraints on complementary specialization to form ecosystems. Combining metabolic universality with species-level plasticity and ecological constraint should clarify the roles that emerging individuality actually played in the origin of life.

References

Borenstein E, Kupiec M, Feldman MW, Ruppin E (2008) Large-scale reconstruction and phylogenetic analysis of metabolic environments. Proc Nat Acad Sci USA 105:14482–14487

Brazelton WJ, Baross JA (2009) Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. ISME J 1–5

Chia N, Woese CR, Goldenfeld N (2008) A collective mechanism for phase variation in biofilms. Proc Nat Acad Sci USA 105:14597–14602

Copley SD, Smith E, Morowitz HJ (2005) A mechanism for the association of amino acids with their codons and the origin of the genetic code. Proc Nat Acad Sci USA 102:4442–4447

Copley SD, Smith E, Morowitz HJ (2007) The origin of the RNA world: co-evolution of genes and metabolism. Bioorganic Chem 35:430–443

de Duve C (1991) Blueprint for a cell. Patterson, Burlington

Hausner M, Wuertz S (1999) High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis. Appl Environ Microbiol 65:3710–3713

Hutchison CA III, Peterson SN, Gill SR, Cline RT, White O, Fraser CM, Smith HO, Venter JC (1999) Global transposon mutagenesis and a minimal mycoplasma genome. Science 286:2165–2169

Kreimer A, Borenstein E, Gophna U, Ruppin E (2008) The evolution of modularity in bacterial metabolic networks. Proc Nat Acad Sci USA 105:6976–6981

Lengeler JW, Drews G, Schlegel HG (1999) Biology of the prokaryotes. Blackwell Science, New York

Martin W, Russell MJ (2003) On the origin of cells: an hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. Phil Trans Roy Soc London 358B:27–85

Martin W, Russell MJ (2006) On the origin biochemistry at an alkaline hydrothermal vent. Phil Trans Roy Soc B 1–39. doi:10.1098/rstb.2006.1881

Martin W, Baross J, Kelley D, Russell MJ (2008) Hydrothermal vents and the origin of life. Nature Rev Microbiol 6:805–814

Mirkin BG, Fenner TI, Galperin MY, Koonin EV (2003) Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. BMC Evol Biol 3:2–35

Mulkidjanian AY, Galperin MY, Koonin EV (2009) Co-evolution of primordial membranes and membrane proteins. Trends Biochem Sci 34:206–215

Nowak MA, Ohtsuki H (2008) Prevolutionary dynamics and the origin of evolution. Proc Nat Acad Sci USA 105:14924–14927

Pereto J, Lopez-Garcia P, Moreira D (2004) Ancestral lipid biosynthesis and early membrane evolution. Trends Biochem Sci 29:471–477

Pohorille A, Schweighofer K, Wilson MA (2005) The origin and early evolution of membrane channels. Astrobiology 5:1–17

Schrenk MO, Kelley DS, Bolton S, Baross JA (2004) Low archaeal diversity linked to subseafloor geochemical processes at the lost city hydrothermal field, Mid-Atlantic ridge. Environmental Microbiology 6:1086–1095

Simon HA (1973) The organization of complex systems. In: Pattee HH (ed) Hierarchy theory: the challenge of complex systems. Braziller, New York, pp 3–27

Sinanoglu O (1975) Theory of chemical reaction networks. all possible mechanisms or synthetic pathways with given number of reaction steps or species. J Amer Chem Soc 97:2309–2320

Sinanoglu O (1984) On the algebraic construction of chemistry from quantum mechanics. a fundamental valency vector field defined on the euclidean 3-space and its relation to the hilbert space. Theoret Chim Acta 65:243–248

Smith E, Morowitz HJ (2004) Universality in intermediary metabolism. Proc Nat Acad Sci USA 101:13168–13173

Srinivasan V, Morowitz HJ (2009a) The canonical network of autotrophic intermediary metabolism: minimal metabolome of a reductive chemoautotroph. Biol Bulletin 216:126–130

Srinivasan V, Morowitz HJ (2009b) Analysis of the intermediary metabolism of a reductive chemoautotroph. Biol Bulletin 217:222–232

Vetsigian K, Woese CR, Goldenfeld N (2006) Collective evolution and the genetic code. Proc Nat Acad Sci USA 103:10696–10701

Woese CR (1998) The universal ancestor. Proc Nat Acad Sci USA 95:6854–6859

Woese CR (2000) Interpreting the universal phylogenetic tree. Proc Nat Acad Sci USA 97:8392–8396

Woese CR (2002) On the evolution of cells. Proc Nat Acad Sci USA 99:8742–8747

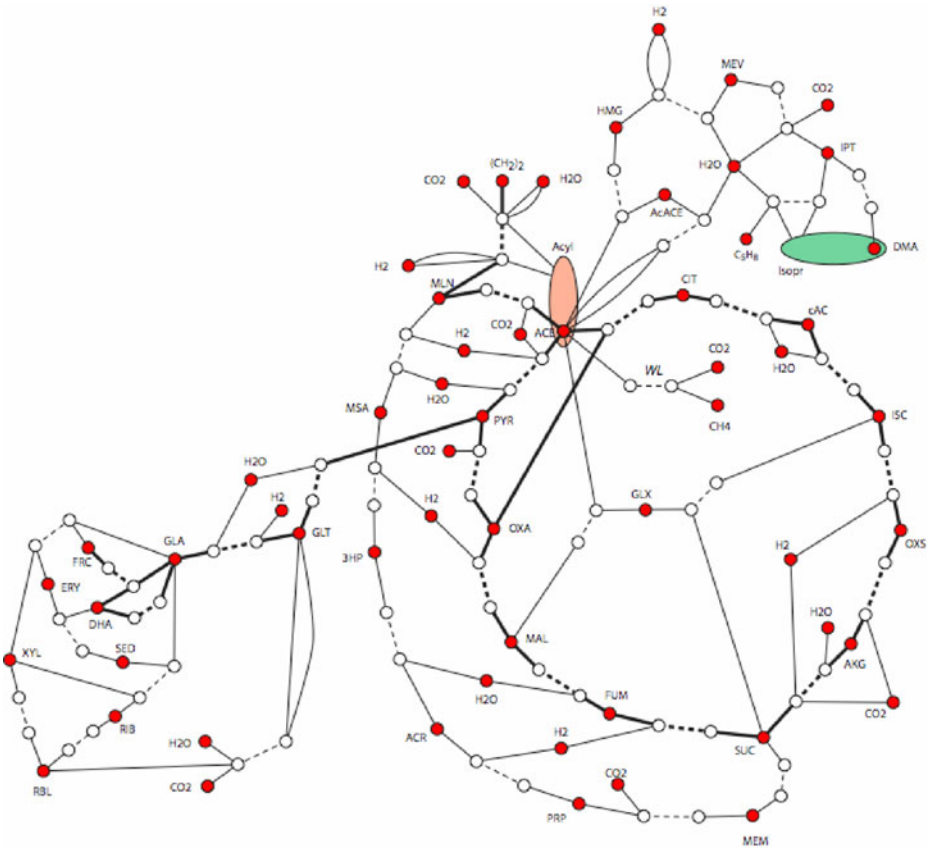


Fig. 1 The universal core of carbon biosynthesis (heavy lines) and major variations. Solid circles represent chemical reagents, and dashed lines represent chemical reactions, in a diagrammatic similar to that proposed by Sinanoglu (1975, 1984). Shaded large ellipses indicate recursively defined pathways: malonate decarboxylation to extend fatty acid chains by hydrocarbon (CH_2) groups, and isoprene (C_5H_8) condensation. Citric-acid cycle, Wood-Ljungdahl, Calvin-Benson, and 3-hydroxypropionate pathways shown, with phosphate and sulfur chemistry omitted and only CHO backbones indicated. Abbreviations: ACE, acetate;

PYR, pyruvate; OXA, oxaloacetate; MAL, malate; FUM, fumarate; SUC, succinate; AKG, α-ketoglutarate or 2-oxoglutarate; OXS, oxalosuccinate; ISC, isocitrate; cAC, cis-aconintate; CIT, citrate; MLN, malonate; AcACE, acetoacetate; HMG, hydroxy-methyl glutarate; MEV, mevalonate; IPT, isopentenyl; DMA, di-methylallyl; GLT, glycerate; GLA, glyceraldehyde; DHA, di-hydroxyacetone; FRC, fructose; ERY, erythrose; SED, sedoheptulose; XYL, xylulose; RBL, ribulose; RIB, ribose; MSA, malonate semialdehyde; 3HP, 3-hydroxypropionate; ACR, acrylate; PRP, propionate; MEM, methyl-malonate; GLX, glyoxylate.

E-mail: desmith@santafe.edu

Modularity and Function in Early Prebiotic Evolution

Alvaro Moreno

Dept. of Logic and Philosophy of Science, Univ. Basque Country, Spain

Keywords Modularity • Self-Assembly • Holism • Functional Parts • Pre-Darwinian Evolution

It is widely accepted that something with the degree of complexity of the simplest present-day cells cannot suddenly appear. Thus, it is assumed that a more or less long *cumulative process* has been necessary. However, it is hard to explain this cumulative process, because neither self-organization nor Darwinian evolution by Natural Selection explain satisfactorily early prebiotic evolution. NS starts with a system able to reproduce genetically, whereas self-organization cannot ensure an iterative accumulation of organizational complexity. So, how can we explain the primitive cumulative process of complexification, from chemical evolution to systems able to evolve by Natural Selection? Following Herbert Simon (1962), recently Fox Keller (2007) has suggested that if stable heterogeneous systems, initially quite simple, merge into composite systems that are themselves—mechanically, thermodynamically, chemically—stable, such composite systems in turn would provide the building blocks for further construction, generating a process in which novelty arises through composition (or combination) and finally becomes integrated into the changing population by selection for increased relative stability.

This idea of evolution by composition supposes a modular-based type of organization. Now, the idea of evolution by composition implies a modular type of organization. What initially were independent systems tend to become, first, inter-dependent modules and later, after an evolutionary process, components of highly integrated networks. The global integrated system gains stability and functional efficiency. For example, according to this view, present-day cellular organization would be the result of an intensive horizontal exchange of genes that explored different combinations of functional modules (Woese 2002). However, modularity is much harder to see when we are considering primitive self-maintaining chemical networks, because (presumably) they were highly distributed: dynamical properties of the global organization ensure the cyclic reproduction of certain local interactions, and the organizational stability of these systems.

An evolution based on accumulation of "modules" is relatively easy to conceive when we consider processes of assembling of thermodynamically conservative structures. Simple building blocks generate spontaneously composite stable structures (atoms, molecules, macromolecules…) due to different levels of forces. Thus, in the abiotic world, assembling

processes—and specially, self-assembly—yield much more structural complexity and modular variety than self-organizing processes. The problem is that this "evolution by composition" is bounded to thermodynamic equilibrium, and living systems are organized matter in far-from-equilibrium (FFE) conditions. Accordingly, we are interested in a concept of modularity within an organizational FFE frame. Now, when we consider modularity in this context, we are confronted with the fact that behind this concept of modularity is that of *functional* parts.

But in primitive FFE chemical networks it is difficult to say in what sense a given (type of) component does something functionally different from another. To harbor some form of functional diversity some parts (or processes) within a system must be capable to perform operationally distinguishable contributions to the self-maintenance/survival of this system. We need to be able to tell how each part specifically affects the global functioning of the system.

I will suggest that an early form of functional diversity can be found in a special type of self-maintaining organization, arising from the interplay among a set of different endogenously produced constraints (pre-enzymatic catalysts and primitive compartments included). It is therefore the result of an association between processes of assembling of thermodynamically conservative structures and distributed, holistic dissipative organizations, as shown for example in a recent model by Ruiz Mirazo & Mavelli (2007). The model simulates real membrane processes coupled to chemical autocatalytic reactions. The "mechanical" dynamics of the membrane is operationally coupled to the chemical dynamics of the autocatalytic network in the following way: when the osmotic pressure attains certain threshold, peptides in the membrane open channels; and this happens because, due to elastic tension (a mechanical process), polypeptides inserted in the membrane adopt the suitable conformation to form waste-transport channels. Thus, these proto-cells show a form of passive but *mediated transport*.

Likely, this type of systems had its origin in formerly independent systems (self-maintaining chemical networks, self-assembling vesicles), each with its own recognizable dynamics and requirements for stability, and when they coupled together they got transformed, becoming strongly inter-dependent (Mavelli & Ruiz-Mirazo 2007; Piedrafita et al. 2009). The model shows that, once integrated, the global viability of the system is the consequence not only of the self–maintenance of a dissipative chemical network but of the interplay between chemical reactions and other types of processes (e.g., self-assembly, diffusion, transport through the membrane). On the one hand, physical changes affect the functioning of the chemical network; but on the other hand, changes in the dynamics of the chemical network can also affect assembling processes. In other words, the way these systems operate is not just a question of 'pure chemistry' (in the sense of abstract reaction pathways or cyclic networks) but chemistry inherently coupled to various biophysical—even mechanical—processes (actually, chemical self-maintaining processes provide the constraints, which harness law-like interactive processes between the building blocks generating certain kind of mechanical structures or arrangements, that in turn will modulate the chemical network in a functional way, namely, improving its stability). In this case it is justified to speak in terms of functional parts because there are very different processes (self-assembly and self-organization) contributing together to the maintenance of the system.

However, we cannot describe this minimal form of functional differentiation as "modules" that can be re-arranged through processes of cut and paste. A modular organization is one in which different functional components may be separated and recombined. Now, this implies that the functional components must keep at least a core of

specific functionality. It is only when the functional parts of a system 1) acquire a degree of stability such that they can persist outside of a (highly) specific form of organization that they become potentially capable of acting as modules; and, 2) they convey a core structure such that, through processes of arrangement of similar structures, an open world of functions can be achieved. In other words, when new functions can be generated through compositional processes of arrangement of similar core structures.

What is lacking in the former example (as in other similar models of proto-metabolic cellular systems, like Ganti's Chemoton (1975, 2003) is precisely this condition. For, despite the fact that we can distinguish certain thermodynamically conservative structures performing functional roles, the suitable arrangement of parts constituting the entire functional *mechanism* is maintained and regenerated by the very organization of the system. Thus, in this system the functional components do not keep a functional core out of this specific organization. Instead, in a modular organization the functional units must be a) relatively stable entities capable of playing minimally similar roles in a set of different organizations; b) they must be amenable to a variety of physical forms of assembly or aggregation; and, c) they must play a rather generic, not highly specific type of functionality. Otherwise, they could not act as building blocks for the construction of new functional entities.

It is not easy to tell what degree of molecular complexity is required to satisfy these conditions. Lacking full-fledged regulatory controls, proto-metabolic networks could hardly show *dynamically based (organizational)* modules (satisfying the former conditions). However, it is conceivable that relatively small molecules could act as structural modules in the sense envisaged by Simon, provided they perform distinguishable functions in a FFE, chemical SM organization and that the new-composed structures can be recruited to perform new functions. Actually, Manrubia & Briones (2007) and Briones et al. (2009) have shown that certain small molecules of RNA can play the role of modules in a stepwise process of ligation-based modular evolution: RNA hairpin modules could have displayed ligase activity, catalyzing the assembly of larger, eventually functional RNA molecules. These ligation processes allow a fraction of the population to retain their previous modular structure, and thus, structural and functional complexity can progressively increase, even in the absence of template replication.

In sum, the origin of life seems to be a process of progressive integration of a former "confederacy". Presumably, these processes of confederation have followed two different stages, an early pre-modular one, and another fully modular later. However, the early form of "confederacy" was quite different of the "Evolution by Composition" proposed by Simon and Fox Keller, which in its spirit requires a clear modular-based form of organization. One of the main advantages of modular-based systems is their capacity for exploring functional novelties. Thanks to the variety of ways/combinations in which functional components may contribute to the maintenance and reproductive success of the systems they belong to, the creation of modular organizations provides a minimal but wide enough phenotypic space for Natural Selection to be actually selective as an evolutionary mechanism.

References

Briones C, Stich M, Manrubia SC (2009) The dawn of the RNA world: toward functional complexity through ligation of random RNA oligomers. RNA 15(5):743–749

Ganti T (1975) Organization of chemical reactions into dividing and metabolizing units: the chemotons. BioSystems 7:15–21

Ganti T (2003) The principles of life. Oxford UP, New York
Keller EF (2007) The disappearance of function from 'self-organizing systems'. In: Boogerd F, Bruggeman F, Hofmeyr JH, Westerhoff HV (eds) Systems biology. Philosophical foundations. Elsevier, Amsterdam, pp 303–318
Manrubia S, Briones C (2007) Modular evolution and increase of functional complexity in replicating RNA molecules. RNA 13(1):97–107
Mavelli F, Ruiz Mirazo K (2007) Stochastic simulations of minimal self-reproducing cellular systems. Phil Trans R Soc B 362:1789–1802
Piedrafita G, Mavelli F, Morán F, Ruiz-Mirazo K (2009) On the transition from prebiotic to proto-biological membranes: from 'self-assembly' to 'self-production'. Proceedings of ECAL'09, accepted for publication (Springer, LNCS Series)
Ruiz-Mirazo K, Mavelli F (2007) Simulation model for functionalized vesicles: lipid-peptide integration in minimal protocells. In: Almeida e Costa F et al (eds) ECAL 2007, Lisbon, Portugal, September 10–14, 2007, Proceedings. Lecture Notes in Computer Science 4648. Springer, Heidelberg, pp 32–41
Simon H (1962) The architecture of complexity. Proc Am Philos Soc 106:467–482
Woese C (2002) On the evolution of cells. PNAS 99(13):8742–8747

E-mail: alvaro.moreno@ehu.es

Commentary on "Life as a Unity or Confederacy"

Addy Pross

Department of Chemistry, Ben-Gurion University of the Negev, Israel

Keywords RNA-World • Metabolic Cycles • Dynamic Kinetic Stability • Origin of Life

The question whether life is a unity or confederacy goes to the very heart of the Origin of Life debate. In essence it asks whether life's emergence began with some unique physicochemical event followed by an extended process of complexification (whose details are yet to be clarified), or life began with the emergence of two or more identifiable subsystems which then merged into the holistic entity of some early life form. The unitary hypothesis is exemplified by the RNA-world view (Gesteland et al. 1999; Joyce 2002; Orgel 1998), which postulates the fortuitous formation of some replicating entity, RNA or RNA-like, which then underwent a process of complexification leading to the emergence of some minimal living system with its associated subsystems. The confederacy view is exemplified by Dyson's "double origin hypothesis" (Dyson 1985), where two subsystems—an independently formed replicating entity and an independently formed metabolic entity—combined to create a complex system, exhibiting both replicative and metabolic capabilities. The question is of practical importance since current attempts to synthesize a minimal living system (reviewed in Luisi 2006) will clearly benefit from an understanding of the path originally taken by nature to achieve that goal. In this commentary we wish to provide arguments favoring the unity view over the confederacy view and note that the above question connects directly with the replication-first—metabolism-first dichotomy (Shapiro 2006; Kauffman 2000; Segre et al. 2000; Morowitz et al. 2000; Orgel 2008; 2000, 1992; Pross 2004; Lazcano and Miller 1999; Lifson 1997; Maynard Smith and Szathmary 1995).

1. *Application of Occam's Razor—"Entities should not be multiplied unnecessarily".*

If, according to the confederacy view, we consider metabolic and replicative capabilities as exemplifying two key subsystems that would have needed to emerge and combine to generate a living system, application of Occam's Razor to that view suggests it to be less likely. Let us explain. Both replication-first and metabolism-first schools of thought remain controversial and subject to on-going polemic debate due to the fact that persuasive chemical arguments against both approaches have been raised. Accordingly we would argue that the requirement for the ***independent*** emergence of each of these two quite special capabilities—replication and metabolism—necessarily weakens the confederacy argument. A theory which is based on the validity of ***one*** uncertain premise seems more desirable (or at least less undesirable) than one that is based on the validity of ***two*** uncertain premises. So, at least with regard the specific merging of metabolic and replicative subsystems, the confederacy viewpoint takes on the combined deficiencies of the two competing schools. However a more general problem with the confederacy approach needs to be considered, one associated with the thermodynamics of aggregation processes.

2. *Kinetic and Thermodynamic Considerations*

If we build an origin of life hypothesis on a confederacy viewpoint, then a critical issue that needs to be explained is the conversion of simpler equilibrium (or strictly-speaking, pseudo-equilibrium) subsystems into the far-from-equilibrium holistic systems that constitute the simplest living beings, e.g., a bacterial cell. Thus, even if the independent emergence of the particular subsystems is accepted (despite the difficulties mentioned above), it remains far from clear how the amalgamation of two or more equilibrium (or pseudo-equilibrium) systems can result in the formation of a far-from-equilibrium composite. Thermodynamic considerations suggest that a physical merging of the three key life subsystems (i.e., metabolic, replicative and compartmental) may succeed in creating a cell, morphologically speaking, but that cell would likely find itself in a pseudo-equilibrium state, that is it would be ***dead***. In chemical terms that system, by the very conditions by which it was created, would not constitute the dynamic far-from-equilibrium state associated with a living cell.

The physicochemical (as opposed to the historical) dilemma in understanding the origin of life requires us to come up with a mechanism by which established physicochemical principles would explain the ***in-principle*** conversion of a ***simple equilibrium (or pseudo-equilibrium)*** system into a ***complex far-from-equilibrium*** system. The unity approach appears to us to address this question more satisfactorily in that we have recently argued that a simple non-metabolic replicating entity could be expected to be transformed by kinetic selection into a metabolic (in the energy-gathering sense) replicating system, and this key transition could be viewed as a first step toward the generation of a far-from-equilibrium replicating system (Pross 2008, 2004). Thus we would argue that a unity view, through a kinetic analysis of replicating systems, may be better able to provide a causal explanation for the spontaneous emergence of (kinetically stable but thermodynamically unstable) metabolic far-from-equilibrium dynamic replicating systems.

References

Dyson F (1985) Origins of life. Cambridge U.P., Cambridge

Gesteland RF, Cech TR, Atkins JF (eds) (1999) The RNA world. Cold Spring Harbor Laboratory Press, NY

Joyce GF (2002) The antiquity of RNA-based evolution. Nature 418:214–221
Kauffman SA (2000) Investigations. Oxford U.P., Oxford
Lazcano A, Miller SL (1999) On the origin of metabolic pathways. J Mol Evol 49: 424–431
Lifson S (1997) On the crucial stages in the origin of animate matter. J Mol Evol 44:1–8
Luisi PL (2006) The emergence of life. From chemical origins to synthetic biology. Cambridge U.P., Cambridge
Maynard Smith J, Szathmary E (1995) The major transitions in evolution. Freeman, Oxford
Morowitz H, Kostelnik JD, Yang J, Cody GD (2000) The origin of intermediary metabolism. Proc Nat Acad Sci USA 97:7704–7708
Orgel LE (2008) The implausibility of metabolic cycles on the prebiotic earth. PLoS Biol 6:5–13
Orgel LE (2000) Self-organizing biochemical cycles. Proc Nat Acad Sci USA 97:12503–12507
Orgel LE (1992) Molecular replication. Nature 358:203–209s
Orgel LE (1998) The origin of life—a review of facts and speculations. Trends Biochem Science 23:491–495
Pross A (2008) How can a chemical system act purposefully? Bridging between life and non-life. J Phys Org Chem 21:724–730
Pross A (2004) Causation and the origin of life. Metabolism or replication first? Orig Life Evol Biosph 34:307–321
Segre D, Ben-Eli D, Lancet D (2000) Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Nat Acad Sci USA 97: 4112–4117
Shapiro R (2006) Small molecule interactions were central to the origin of life. Quart Rev Biol 81:105–125

E-mail: addy.pross@gmail.com

Biological Autonomy and Systemic Integration

Leonardo Bich

Post-Graduate School of Anthropology and Epistemology of Complexity, University of Bergamo, Italy

Keywords Biological Autonomy • Organization • Autopoiesis • Chemoton Theory

Introduction

The approaches to the study of the origins of life—and to the characterization of the minimal living system as well—based on the hypotheses of unity or confederacy have different implications from the theoretical, epistemological and heuristic points of views. From a theoretical standpoint the fact of considering how strictly subsystems are integrated in a living organism can be useful in order to understand more in depth what makes an organism a system of a certain class. Also, it allows us to deal with the problem of how being part of that system constrains or influences the behavior of the individual

components, that is: to face the problem of downward causation in biological systems. An answer to this question has therefore important consequences as it can open the way to a shift of point of view in the investigation of the living from the properties of specific molecular components—from which to reconstruct life—to the properties of the system they realize, that is, to the conditions the constituents must satisfy in order to be part of it.

From the epistemological point of view, facing this dichotomy means to consider which level of description is, at least in principle, more pertinent in order to catch the specificity of the system under study: that concerning its basic constituents or an higher one. Heuristically speaking this distinction furthermore implies different modalities of fractionation of systems and of identification of relevant components, as I will show in the following.

Two possible "middle way" solutions

I think that a preliminary step in order to answer this question should consist in avoiding the radical dichotomy between two extreme positions on this issue: the mechanistic decomposition of the organism due to a strong version of modularity (Simon 1973), and the holistic view of living systems as not analyzable wholes: the first because it cannot catch the difference between a living system and a machine, the second one because it does not allow any fractionation of the system and, thus, any description of its internal dynamics.

This step brings us to a "middle-way" approach to the problem. It is focused on the role of organization, that is, on the way components are related in order to realize the organism. It aims at providing a description of the autonomy which characterizes biological systems in terms of their generative dynamics, which can be defined through the properties of self-production, self-maintenance and self-distinction from the environment.

In order to follow this pathway of investigation it is necessary to assume a meta-level of description focused on this global internal generative dynamics. From the analysis of the different possible "meta-structures" (Minati 2008)—material or functional—that we can identify at this level two distinct sub-poles of the dichotomy emerge. Both derive from coherent frameworks and they differ in the point of view and in the emphasis given to confederative or unitary aspects. As a consequence they somewhat diverge in the character of their implications.

The first pole consists in a thesis on the partial decomposability of living systems into coupled but semi-independent subsystems. An example of it is provided by Tibor Ganti's Chemoton theory (Ganti 2003). The second one is based on a more strict interdependency of sub-processes which can only be characterized in the light of the higher level system they integrate. This line of thought has been brought forth by the tradition of studies on biological autonomy (Maturana and Varela 1980; Rosen 1991).

The approach focused on the semi-independence of subsystems, exemplified by Chemoton theory, consists in an attempt to answer to the issue of the basic characterization of the living through the identification of a list of the components which are necessary for the realization of the generative dynamics defined before. Ganti indentifies three coupled subsystems:

a) a chemical motor (an autocatalytic metabolic subsystem)
b) a chemical information systems (the control subsystem)
c) a chemical boundary subsystem (the membrane subsystem)

According to this framework the basic living system is characterized by a topological *closure*, due to the production of a membrane, and by a specification of components

instructively induced by a control subsystem. Its generative dynamics is characterized by the coupling between its subsystems according to a model of co-evolution between interrelated semi-independent entities. These are identified according to their intrinsic properties following a bottom-up observational direction. The organization of the system can be therefore defined as "structural", because the relational properties of the constituents can be derived from their structural intrinsic ones.

Autopoietic theory is an example of the second approach, characterized by the thesis of interdependency of sub-processes. It is based on a systemic assumption: the characterization of the living is primary focused on the global organization rather than on the properties of the material constituents or subsystems (Maturana and Varela 1980)[2]. It consists in a reinterpretation of the cybernetic notion of circular self-stabilization which is not to be applied to single regulatory processes or subsystems, then coupled together, but to the whole living system (Bich and Damiano 2008). What is proposed here is a second order cybernetic loop of realization and conservation of the unitary organization of the organism (Bich 2008; Cornish-Bowden and Cárdenas 2008). According to this model the living system is characterized by both a topological and functional *closure*, such that the interdependency of subsystems derives from the higher order organization which defines the identity of the whole system. In this case the relevant components are not the material ones, specified by an instructive subsystem, but functional sub-processes, which can be identified according to their relational properties and with respect to the system they integrate.

By comparing the two "middle-way" approaches we can see how in the first model, more focused on the properties of components, the meta-level of description is primarily structural, or at least the relational aspects can be reduced to the structural ones. In the second one, whose point of view is mainly focused on the global invariance, the structural and relational aspects are instead irreducible and complementary. According to these remarks, I sustain the thesis that it is the presence or not of the higher order integrative circularity—the autopoietic organization—what marks the difference between the two perspectives. In fact it provides the theoretical explanation of the looseness or tightness which the coupling between subsystems can assume.

From the epistemological point of view the opposition between unity and confederacy entails a further opposition between two distinct observational operations: the "structural" and the "functional" identification of subsystems (Bich 2009). In the former—"confederative"—approach, the subsystems are partially independent and can be identified logically and phenomenologically *ex ante* with respect to the realization of the living system they belong to (bottom-up approach): they are the material parts of the systems, characterized through their intrinsic properties. In the latter they can be characterized and identified only *ex-post* and with respect to the unity they integrate. Their condition of existence, in fact, is the presence of the biological system they realize through their interaction (top down approach): they are characterized through their relational properties. In this second approach the identification of the relevant components is more problematic, as this

[2] The autopoietic organization can be defined as a net of processes of production, transformation and destruction of components, that: (1) through their interactions and transformations recursively realize and regenerate the same network that produces them; and (2) constitute the system as a concrete unity in the space in which they exist, by establishing its boundary and thus specifying the topological domain of its realization (Maturana and Varela 1980).

operation requires a multilevel complex strategy which needs to take into consideration both structural and relational aspects in order to provide a description of the generative dynamics of the basic living system.

Concluding remarks: two ways of investigating the origins of life

The frameworks outlined here besides providing distinct models of the basic living systems, entail also two different strategies of facing the problem of the origins of life.

In the perspective based on the hypothesis of the semi-independence of subsystems the main strategy consists in identifying the pertinent basic components and in considering the process of their reciprocal stabilization inside a topological boundary. From both the theoretical and historical points of view, in fact, the emphasis is put on the role of assemblative and coupling aspects in the constitution of the living. Instead, in the autopoietic framework, characterized by a stronger interdependency of sub-processes, this strategy is only the first part of a more complex procedure which moves on different descriptive levels in order to combine the bottom up and the top down approaches. According to this approach the further operations would consist in a relational analysis of the functional components of the global organization and in an attempt to trace back the historical dynamics of the divergence between structural and relational aspects in early evolution.

In a wider perspective, in fact, these two strategies involve not just different models of the living but also different views of natural history. In the first case the implicit idea is that of a continuous process of complexification through the combination of some initial constituents and the stabilization of their interactions. In the second approach the implicit model of natural evolution is closer to Whitehead's idea of process (Whitehead 1929) and it consists in a discontinuous creative process: a succession of emergences in which at any new step we assist to new reorganizations which give rise to new relational unities (Bocchi and Ceruti 1993; Bich 2008; Kauffman 2008).

References

Bich L (2008) L'ordine invisibile. Ripensare il vivente oltre la metafora computazionale. Ph.D. dissertation, University of Bergamo

Bich L (2009) Downward Causation and Relatedness in Emergent Systems: Epistemological Remarks. In: Minati G, Pessa E, Abram M (eds) Processes of emergence of systems and systemic properties. Towards a general theory of emergence, World Scientific, Singapore, pp 591–601

Bich L, Damiano L (2008) Order in the nothing: Autopoiesis and the Organizational Characterization of the Living. EJTP 16/1:343–373

Bocchi G, Ceruti M (1993) Origini di storie. Feltrinelli, Milano

Cornish-Bowden A, Cárdenas M L (2008) Self-Organization at the origin of Life. J Theor Biol 252:411–418

Ganti T (2003) The Principles of Life. Oxford University Press, Oxford

Kauffman S (2008) Reinventing the Sacred. Basic Books, New York

Maturana H, Varela F (1980) Autopoiesis and Cognition. Reidel Publishing, Dordrecht

Minati G (2008) New Approaches for Modelling Emergence of Collective Phenomena. The Meta-structures project. Polimetrica, Milano

Rosen R (1991) Life Itself: a Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life. Columbia University Press, New York

Simon H (1973) The Organization of Complex Systems. In: Pattee H (ed) Hierarchy Theory, Braziller, New York, pp 1–28
Whitehead AN (1929) The Process and Reality. MacMillan, London

E-mail: leobich@libero.it

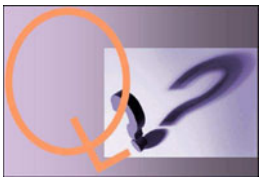
Workshop OQOL'09

Extended Abstracts for the following selected question

- **Defining the very origin of life**

Premise. Defining life in an universal way is notoriously a difficult or impossible task, but also the notion of "origin of life" appears to be rather confuse. Some authors talk about origin of life at the level of the origin of low molecular weight compounds, obtainable either through hypothermal vents; or the pyrite reaction; or by Miller's type of processes. However, you can have all low molecular weight compounds of this world, and you will never be able to make life, as life only arises at the level of specific macromolecular sequences like enzymes, DNA, RNA.

The question. Do you agree that we should have a critical review of the terminology of "origin of life", and, for example, not use this term at the level of low mol. weight compounds (where we have "prebiotic chemistry", or origin of reductive power…), and restrict it instead to the level of the biogenesis of specific macromolecules and their interactions?



Replication and Darwinian Selection Define Life's Origin

Omer Markovitch, Aron Inger, Barak Shenhav, Doron Lancet

Dept. Molecular Genetics, Weizmann Institute, Israel

Keywords Early Evolution • Lipid World • Compositional Information • Mutually Catalytic Network • Stationary State • Non-covalent Assembly • Replication Without RNA

Darwinian process as prerequisite

There are diverse opinions regarding the definition of life's origin, and it is often said that such definition is crucially dependent on how one defines life. We suggest here a stand-alone definition for the origin of life, irrespective of how "life" is

defined. Such definition of origin rests upon the most important foundation of biology—Darwin's theory of natural selection and evolution. We claim that life originates upon the spontaneous emergence of the first chemical entity capable of a Darwinian process.

It is widely accepted that the most crucial element of Darwinism is self-replication with variations. As a side-line, it should be stressed that many authors argue on whether the term "replication" only applies to the copying of linearly coded information, e.g. in polynucleotides. Often this is extended to a claim that no life can emerge without the involvement of "digital" polymeric information transfer. In contrast, we use here a chemistry- and pathway-independent terminology whereby "replication" entails the copying, from one generation to another, of information of any kind, with any measurable level of fidelity. Consequently, we propose that life begins at the inception of a chemical entity, whose internal kinetic and thermodynamic mechanisms result in the appearance of its own simile(s).

It is likely that at the very early (prebiotic) steps of life's inception, replicative processes of discrete molecular entities would be highly inefficient. Thus mutations would likely come naturally, and diversity of progeny would constitute a highly expected outcome. Thus, only one of the above mentioned prerequisites for a Darwinian process, namely replication, is problematic, and necessitates scrutiny. Once replication occurs, variation would ensue, and combined with natural selection, it would result in Darwinian evolution.

Molecular assemblies beget progeny

A major claim to be made here is that life begins at the moment a prebiotic molecular entity acquires sufficiently intricate internal chemistry to allow replication to occur. Our line of thought is different from many other relevant treatises, which address a single self-replicating molecule, typically RNA (Orgel 2004). Here we consider a replicating assembly of N_0 molecules which leads to the formation of two progeny assemblies of the same size. This is a broader definition, with no loss of generality, since one could always choose $N_0=1$. We do not find it necessary to specify the exact chemical nature of the molecules that constitute the assembly, as long as they have general properties (e.g. amphiphilicity) that make them adhere to each other non-covalently and non-specifically in an aqueous medium.

A point of strength of the foregoing definition is that it is strongly independent of specific chemical mechanisms, such as linear covalent polymerization and templating via base pairing. Instead, suffices to assume a set of chemical reactions within the assembly that result in the generation of a second similar assembly. To assist in convincing the reader that such a spontaneous chemical pathway from a single assembly to two similar assemblies is feasible under the laws of chemistry, we resort to our published work on the Graded Autocatalysis Replication Domain (GARD) model, as detailed below.

Molding of environmental chemicals

Practically every process of prebiotic replication that has ever been considered involves the molding of externally supplied molecules. In the textbook scenario of an RNA molecule capable of instructing its own copies, the externally supplied molecules are the ribonucleotides. The crucial process that needs to be addressed, therefore, is how, for a given proposed mechanism, externally supplied building blocks are coerced to construct the new copy of the originally existing entity. A case in point is the

organization of mononucleotides along an existing single-stranded RNA polymer to generate an antisense strand. This is mediated by non-covalent base complementarity of the externally supplied mononucleotides, followed by covalent stringing of such monomers.

When asking how an assembly of non-covalently accreted molecules can undergo replication, we actually face is a question about the influence of the assembly's components on the joining of external molecules. Perhaps the simplest relevant mode of action would be the growth of an assembly, upon absorbing more molecular components, in a way that will preserve the relative proportions of its components (homeostatic growth: Segré and Lancet 2000; Shenhav et al. 2005). While this in itself does not lend itself to the straightforward definition of replication, it is easy to be convinced that if such a process actually happened, only one relatively simple step is necessary to complete a cycle of replication. This step is fission, whereby the larger assembly emerging from a growth process splits into two rather similar halves. When this happens, the accretion from the environment of additional copies of each and every molecule type, making up the original assembly, is translated into the generation of two mutually similar progeny. This unorthodox mode of self-copying, which involves no covalent bond formation or disruption, and definitely no linear molecular templating, still answers the simple definition of simile production, hence can serve as a basis for defining a Darwinian origin of life.

Compositional information and data analysis

Progress in the definition of replication in the context of life's origin has, in our opinion, been hampered by adhering to definitions based on digital information.

True, the notion that RNA or RNA-like polymers were at the core of life ever since the first replicator emerged is attractive and elegant. One of the strong arguments for an RNA world scenario is the capacity of polynucleotides to store and transmit information in a highly defined manner (Orgel 2004). This is based on linear sequences composed of monomers derived from a specific biochemical alphabet. Sequence based information can then be translated to the protein level, thus affording the emergence of elaborate and highly specific catalysts (polypeptides).

However, it is rather clear that digital information, of the kind embodied in DNA and RNA rests on highly elaborate chemistry, including the long biopolymers themselves, as well as the intricate protein or RNA catalysts necessary for their copying. Such mechanisms are not easily envisaged as arising at the very early stages of prebiotic evolution. Taking this into account this, as well as additional crucial points of criticism directed against an RNA-first scenario (Shapiro 2006), we and others have proposed that to consider alternative scenarios. The most injurious hurdle in considering such alternative is finding a substitute for linear, digital information storage and transmission from one generation to another. What comes to rescue is the notion of compositional information.

Compositional information and compositional data analysis are widespread concepts in many branches of science (González 2007). Compositional data are quantitative descriptions of the parts of an entity, represented by a vector whose elements represent relative amounts of the entity's components. It is easy to realize that a composition carries comparable information to a sequence (Fig. 1). When a molecular assembly replicates by accretion followed by fission, it is compositional information that gets propagated. This becomes a crucial element in the potential acceptance of an early life scenario without nucleic acids.

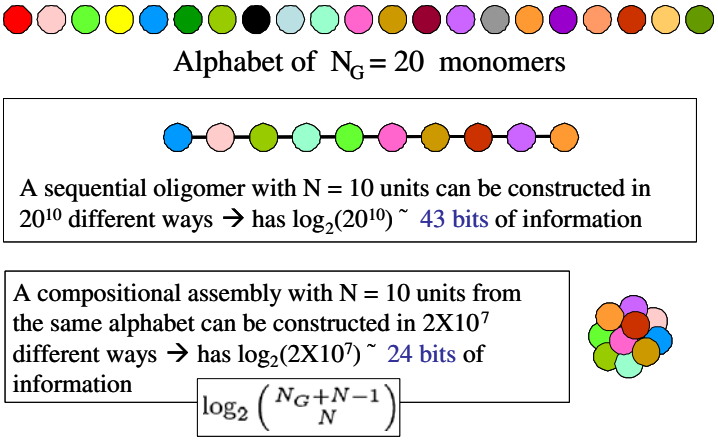


Fig. 1 Comparing compositional to sequence information—an example

The graded autocatalysis replication domain (GARD) model

Our own research on the GARD model provides an example of self-replication without nucleic acids (Segré and Lancet 2000; Shenhav et al. 2005; Shenhav et al. 2007). The entities that generate their own (rather imperfect) copies are assemblies of amphiphiles, and the information being copied from one generation to another is compositional, i.e. the ratios of different types of molecules, independent of spatial arrangement. The mechanism by which such replication can emerge, as shown by our computer simulations, is mutually catalytic networks (Kauffman 1986) that effect molecular accretion and synthesis. A highly important facet of such simulations is the discovery that compositional assemblies may reach (under certain conditions, and for part of the simulation time period), one or more privileged compositional states, termed composomes (Fig. 2), which "breed true" by homeostatic growth and fission. The underlying mutually catalytic networks can be numerically simulated explicitly based on parameters derived from real-world molecular interactions (Shenhav et al. 2005, 2007).

Composomes are capable of undergoing mutation, selection and evolution-like processes (Shenhav et al. 2007; Hunding et al. 2006). The GARD model exemplifies systems that can serve as initiators of prebiotic evolution towards life. If generalized (Shenhav et al. 2005) and further developed, GARD analyses can to help delineate a general definition for life's origin, as well as rigorous criteria for evaluating proposed models for the origin of life. An example of relevant properties of the GARD model is its capacity to emulate the origin of homochirality, i.e. the enantiomeric excess invariably found in all present-day biology. GARD simulations show that such excess is a result of molecular selection dynamics rather than a prerequisite for life's origin (Shenhav et al. 2004).

Conclusion

In summary, we propose that despite its attractiveness, a single molecule sequence-based replication model does not constitute an absolute requirement for defining life's origin. We claim that the appearance of Darwinian selection *is* an essential part of such definition, but that it can be embodied in simpler chemical systems, perhaps such based on mutually catalytic accretion and on compositional information inheritance. The latter are considerably more likely to occur in a spontaneously-emerging chemical entity. Such simpler

systems may show many of the hallmarks of evolution and may lead by subsequent steps of selection, to polymer based life as we know it.

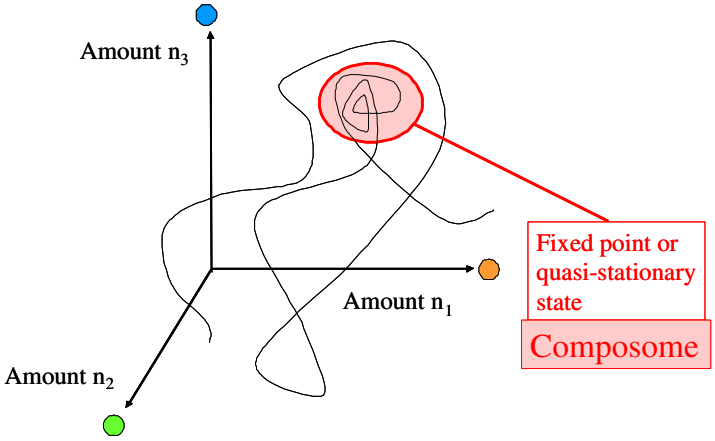


Fig. 2 GARD dynamics—trajectory in a N_G-dimensional compositional space displayed with dimensionality reduced to 3 by Principal Component Analysis (PCA). Dynamically favored region in N_G-dimensional space (resembling fixed points or quasi-stationary states (Dyson 1982)) are composomes

References

Dyson FJ (1982) A model for the origin of life. J Mol Evol 18(5):344–350

González AG (2007) Use and misuse of supervised pattern recognition methods for interpreting compositional data. J Chromatogr A 1158(1–2):215–225

Hunding A, Kepes F, Lancet D, Minsky A, Norris V, Raine D, Sriram K, Root-Bernstein R (2006) Compositional complementarity and prebiotic ecology in the origin of life. Bioessays 28(4):399–412

Kauffman SA (1986) Autocatalytic sets of proteins. J Theor Biol 119(1):1–24

Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. Crit Rev Biochem Mol Biol 39(2):99–123

Segré D, Lancet D (2000) Composing life. EMBO Rep 1(3):217–222

Shapiro R (2006) Small molecule interactions were central to the origin of life. Q Rev Biol 81(2):105–125

Shenhav B, Kafri R, Lancet D (2004) Graded artificial chemistry in restricted boundaries. Proceedings of the 9th Intl Conference on the Simulation and Synthesis of Living Systems (ALIFE9)

Shenhav B, Bar-Even A, Kafri R, Lancet D (2005) Polymer GARD: computer simulation of covalent bond formation in reproducing molecular assemblies. Orig Life Evol Biosph 35 (2):111–133

Shenhav B, Oz A, Lancet D (2007) Coevolution of compositional protocells and their environment. Philos Trans R Soc Lond B Biol Sci 362(1486):1813–1819

E-mail: doron.lancet@weizmann.ac.il

Hierarchical Definitions in the Origin of Life

Jim Cleaves

Carnegie Institution for Science, Washington D.C, USA

Keywords Biological Autonomy • Organization • Autopoiesis • Chemoton Theory

It is notoriously difficult to define "Life" descriptively, even in chemical terms (Schrödinger 1945; Cleland and Chyba 2002; Palyi et al 2002; Ruiz-Mirazo et al. 2004). Defining the origin of life must depend on first having a solid definition of life, which can be distinguished from a Theory of Biology (Stein and Varela 1994), which might describe the essential functioning of life. This problem can be further divided into theoretical definitions (Ganti 1993; Szathmary 2006) of origins, as opposed to operational definitions (Fleischaker 1990). In order to be experimentally useful, a set of criteria which would allow the classification of a chemical system as living is required. Life detection in geological samples (extant or extinct, terrestrial or extraterrestrial), is a related problem since "detection" depends on the presence of certain pre-agreed diagnostic characteristics (Klein et al. 1976; Davis and McKay 1996; Brasier et al. 2002; Schopf and Bottjer 2009).

Leaving aside more exotic possibilities (i.e. silicon-based life), we may build up a hierarchy of necessary but not sufficient characteristics that a living system might need to possess, for example: 1.) be carbon-based (Benner et al. 2004), 2.) be composed of a non-equilibrium set of organic compounds (in terms of type and/or chirality with respect to known abiotic mechanisms of synthesis) (Weber and Miller 1981; McKay 2008), 3.) be capable of self-reproduction (Szathmary 1996; Benner et al. 2004), 4.) be capable of passing on heritable structural and/or functional mutations (Szathmary 1996 et al. 2004) (Figure 1).

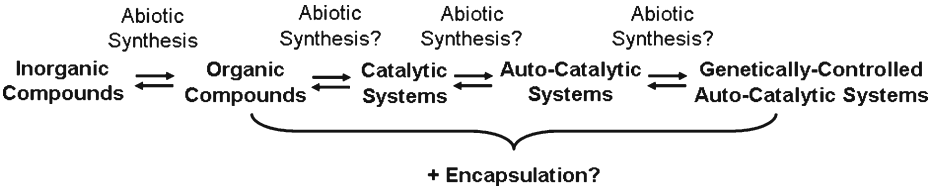


Fig. 1 Possible hierarchical assemblages useful for defining the origin of life

This list is open to debate and not exhaustive, and the sequential nature implied in Figure 1 is not likely entirely necessary. In addition, whether any of these functionally-described systems requires encapsulation is open to discussion. The line between extant and extinct life could be drawn between criteria 2 and 3, while that between prebiotic chemistry and life might blur across 1–3. The failure of a system to meet the required criteria may be useful for refining the criteria.

References

Benner SA, Ricardo A, Carrigan MA (2004) Is there a common chemical model for life in the universe? Curr Opin Chem Biol 8:672–689

Brasier MD, Green OR, Jephcoat AP, Kleppe AK, Van Kranendonk MJ, Lindsay JF, Steele A, Grassineau NV (2002) Questioning the evidence for Earth's oldest fossils. Nature 416:76–81
Cleland CE, Chyba C (2002) Defining 'life'. Orig Life Evol Biosph 32:387–393
Davis WL, McKay CP (1996) Origins of life: a comparison of theories and application to Mars. Orig Life Evol Biosph 26:61–73
Fleischaker GR (1990) Origins of life: an operational definition. Orig Life Evol Biosph 20
Ganti T (2003) The principles of life. Oxford, USA
Klein HP, Lederberg J, Rich A, Horowitz NH, Oyama VL, Levin GV (1976) The viking mission search for life on Mars. Nature 262:24–27
McKay CP (2008) An approach to searching for life on Mars, Europa, and Enceladus. Space Sci Rev 135:49–54
Orgel LE (1998) The origin of life—a review of facts and speculations. Trends Biochem Sci 23:491–495
Palyi G, Zucchi C, Caglioti L (2002) Fundamentals of life. Elsevier
Ruiz-Mirazo K, Peretó J, Moreno A (2004) A universal definition of life: autonomy and open-ended evolution. Orig Life Evol Biosph 34:323–346
Schopf JW, Bottjer DJ (2009) World summit on ancient microscopic fossils. Precambrian Res 173:1–3
Schrödinger E (1945) What is life? The physical aspect of the living cell. Macmillan
Stein W, Varela FJ (1994) Thinking about biology. Santa Fe Institute Studies in the Sciences of Complexity: Lecture Notes. Westview Press
Eörs Szathmáry (2006) The origin of replicators and reproducers. Phil Trans R Soc B 361 (1474):1761–1776
Weber AL, Miller SL (1981) Reasons for the occurrence of the twenty coded protein amino acids. J Mol Evol 17:273–284

E-mail: hjcleaves@ciw.edu

Ranking of Sites on Early Earth as Cradles for Life

Z.P. Zagórski

Department of Radiation Chemistry, Institute of Nuclear Chemistry and Technology, Warsaw, Poland

Keywords High Energy Chemistry • Ionizing Radiation • Panspermia • Prebiotic Chemistry • Radiolysis Location

Proposal of systematics

By analogy to informatics, one can discuss the prebiotic chemistry on primitive Earth as divided into chemical reactions (e.g. those leading to the RNA world, that is the "software" now developed in contemporary laboratory) and into physico-chemical transformations which proceed in real ancient geophysical conditions (e.g. fractional crystallizations, natural "chromatographic" separations on sand dunes watered by oceanic waves, etc). One can also invoke the analogy with chemical technology dealing with its unit processes (a sequence of

chemical reactions) and unit operations (e.g. a small crater filled with water, collecting reactive fallout from the nearby volcano).

Only the terrestrial "hardware" acting as a natural laboratory, integrated with proper chemical "software" could have been functioning as a cradle for Life. All ideas on possible chemical mechanisms leading to prebiotic compositions are developed in laboratory glassware, without reference to real conditions.

There is a sizeable library of chemical "software", supposed to proceed on primitive Earth. Hundreds of chemical reactions expected to run as prebiotic on early Earth are investigated in the laboratory and sometimes in sofisticated installations, like for Fischer-Tropsch synthesis, in complicated arrangements like for ammonia formation on minerals etc. Therefore, not all of them had equal chances to happen. For instance, all catalytic, homogenous reactions including autocatalytic ones, demand rather high concentrations, location in closed containers, all that according to chemical kinetics. Reactions involving heterogeneous catalysis had higher chances, e.g. if the walls of the containment were at the same time catalytically active.

One of the key problems of the hardware in prebiotic chemistry is supply of energy which provided the change of internal energy of the particular system on early Earth. It includes photochemical energy, rather neglected in considerations of creation and development of prebiotic chemical systems. Little is known about the spectrum of the young Sun and filtering properties of the atmospheres, different from the present values, keeping the functioning of life now.

Stanley Miller's experiment as the best example of software+hardware combination

Already the classic case of Stanley Miller's experiment has translated easily into the landscape of primitive Earth, with atmospheric electric discharges supporting energetically the formation of organics, via free radicals and other reactive intermediates. An important difference between chemical reactions realized in the laboratory and the same reactions supposed to run on primitive Earth consists in a low amount of new inorganic products and, more importantly, organics (in relation to large volume of the medium in the real life conditions). In the case of classic Miller reaction, the batches of amino acids are formed rapidly, as sequences of free radical reactions, but they almost disappear in large mass of water containment. That reservoir, as the kind of hardware, could act as a storage container serving for collection of primary products to be used later after ages of storage. The solubility of racemic amino acids, products of Miller reactions, is lower than homochiral components, therefore they form precipitate after reaching well defined concentration, and due to higher density than the mother liquor, they are falling to the bottom of the aqueous container, waiting there, perhaps for other chemical changes and transformations. Neither Miller, nor his students and followers, as well as no other chemists investigating synthesis of new compounds by high energy chemistry in the primitive atmosphere, have tried to speculate what will happen after thousands of cycles of their reaction repeated in millions of years. There is a need for a geological approach to the formation of layers of racemates at the bottom of "Darwinian warm ponds". All sites hosting prebiotic chemistry are important not only because of the possible site of origin of Life. If they involve the place in which organic products formed can be deposited during thousand of years, they can be considered as the producers of the feedstock for microorganisms appearing later. Dried "Millers soup" containing amino acids, even as racemates, usually can be later partly consumed, leaving D-amino acids as waste. Therefore, speculations on the prebiotic chemistry running in the natural "hardware" should be considered not only as the raw material for the origin of Life, but also as the food for

living creatures newly formed. In the "astronomical" scale of time involved, deposits of such "food" can be of substantial size. As concerns other planets: Assuming that prebiotic chemistry was running in a similar way on Mars but stopped half way before origin of life, rich layers of racematic amino acids should be left on the surface after the evaporation of Martian Oceans. There is no sign of it, so either Miller's scenario did not work on Mars, or organics formed were later destroyed by ionizing radiation during billions of years of exposition of naked planet to high energy radiations.

Examples of other "hardware" possible on early Earth and their properties

- Another feature of real chemical conditions on prebiotic Earth are periodically repeated cycles of temperature change due to day/night, more frequent than now, and longer temperature cycles due to seasonal summer/winter variations. These variations can be helpful to explain some physico-chemical separations. For instance periodical partial freezing of aqueous solution can be accompanied by formation of higher concentration of salts and new formed organic compounds. Natural separation of crystals of ice from the solution and repetition of the phenomenon could pay tribute to the lowering of entropy of the system, desired for explanation, from the point of view of the origin of Life.
- The "open air stage" may be not the best site for early Life. The periodicity of the endless repetition of "software" in specific "hardware" conditions, may not contribute to the effective development of prebiotic chemistry as the introduction to life. In that case, Nature offered places on Earth, screened from day/night fluctuations and flattening the seasonal summer/winter effects. These places are caves with underground lakes, rather seldom in prebiotic Earth, because of volcanic origin, formed only sometimes in basalt rocks. Present day, plentiful caves are product of life: calcium carbonate rocks penetrated with streams of water; they did not exist on early Earth.
- More popular sites, isolated from day/night and from seasonal changes of weather, are bottoms of aqueous world. Photochemistry does not exists already on reasonable depths and no temperature effects exist at almost thermostating conditions. No storms are reaching the depths and movements of water are very slow, or nonexistent in isolated lakes.
- There are other possible "hardware", like crystalline, specifically layered minerals at the bottom of aqueous pond, monochiral crystal faces, reactive solids supplied by volcanic eruptions, concentrated sources of ionizing radiations, even aerosols suggested as sites of prebiotic chemical reactions. One can speculate in what way these can be combined with laboratory "software" to create new proposals for origins of Life. In the author's Laboratory a construction is proposed in the frame of enlargement the present Radiation Centre, of a complicated box, with a controlled atmosphere, according to a supposed composition of one of accepted version of early Earth. The chamber, carefully isolated from the rest of the laboratory, will be kept sterile, therefore avoiding difficulties met by Miller who had to add strong bacteria killing compounds to prevent contamination of his soup by present day bacteria and products of their metabolism.
- Special "hardware" conditions have to be considered in the case of photochemical "software". The range of light quanta in the UV/Vis range in the medium with chromophoric groups is short. Depending on the value of the ε

(molar extinction coefficient) and concentration, the range is seldom deeper than 1 mm. The proper example, but in the developed life world, is a plants leave, always thin, absorbing the sun light with production of oxygen from water and of carbohydrates from CO_2. Returning to prebiotic chemistry, pure water exhibits transparency to UV/Vis and therefore light absorbing object, e.g. a micelle can be several centimeters below the surface and will be reached by quanta, if matched to the absorption spectrum.

- Some sites supposed to harbor developing prebiotic chemistry are described in the literature, for instance vents of hot water at the bottom of oceans. They contain rich spectrum of inorganic reactive material which can turn into organics. Some minerals can have catalytic effects; sulphides can, hopefully acts both as a cradle and at the same time as catalyst. However, the present life around the vent shows that there are strictly defined, narrow, thin circles around the orifice, with proper conditions for particular species. Minimal cells, not able to stay at the same place in the moving medium around the vent, could have their niche in layered minerals, securing a stable position in proper place.
- One could discuss further different possibilities of "hardware" in the origin of Life. For instance, one can imagine formation of a shallow meteorite crater, filled with rain water initially and later supplied by small stream, keeping from drying off. If placed near a volcano, the crater would be supplied by volcanic mineral ejecta. In the case of calcium carbide present in it, the reaction with water would yield highly reactive acetylene, starting an efficient chemical "software". The interesting organic C-C-bond occurring unexpectedly from inorganic volcanic magma, could play an important role in further reactions, involving, e.g. secondary organic products of electric discharges in the atmosphere above the crater, as well products of photochemical reactions, in the case of shallow ponds.

Ranking of possible "hardware" in the prebiotic history of the Earth, possibly connected with minimal cell.

The list of "hardware", even simple ones and obvious in the laboratory, but imagined to work on early Earth, is limited by the modest landscape prevailing at that time. Chosen sites on primitive, lifeless Earth, ready to harbor chemical "software" sequences of prebiotic chemistry, most promising in turning into cradles of Life, are now ordered in the ranking from the most probable to the most difficult to be imaged. The list is responsibility of the author only and there are no references, except observations of not published discussions during Conferences on the subject.

- The surface of the Earth covered in substantial degree with oceans of still changing mineral content and a reactive atmosphere above;
- Electric discharges in the atmosphere, of various energy and intensity of single acts, sufficient to produce reactive intermediates;
- Movement of substantial parts of land, creating favorable niches, geochemical effects, producing new minerals and structures, formation of redox ions, e.g. Fe II/III in solution, participating in inorganic/organic reactions, also of chain character;
- UV/Vis light, of the rich spectrum, emitted by the young Sun, generating excited chemical entities in media containing proper chromophoric groups;
- Photochemistry in shallow ponds, continuous exchange of products with deep layers;
- The presence of other sources of energy, feeding the high energy chemistry, like radioactive isotopes of variety of energy of quanta and/or particles and general activity, in some cases by two orders of magnitude higher than now;

- Layered clays, like montmorillonite with expandable nanosized niches for chemical reaction fed by diffusion of reagents and supply of energy. Additional catalytic effects on the walls of the nest;
- Intensive volcanic activity on land, producing variety of compounds, including important carbon compounds in neighboring aqueous ponds;
- Occasional bombardment of Earth by meteorites and asteroids, changing the surface of the Earth and supplying some exotic chemical compounds;
- Volcanic activity at the bottom of seas, creating zones of different temperature and chemistry around vents;
- large differences of temperature from freezing (probably) to the heat close to the boiling point of water, also climatic changes, in the day/night rhythm more rapid than know, seasonal variations summer/winter;
- more intensive tides than now, due to closer distance of the Moon to the Earth, causing changes of the coastline and penetration of sea water into the land.
- All these sites contribute to heterogeneity of chemical changes which might be contributing positively to the origin of Life.

Final comments and Outlook

Unfortunately, the problem of hosting chemical reactions on the early Earth is continuously interfered by occurring time and now the idea of panspermia, supposed to solve all problems of origin of Life on Earth. It has been shown, that radiation chemistry following ionizing radiation in outer space shows impossibility of panspermia, even of transportation of blocks of chemical structures, precursors of Life (Zagorski 2007). Acceptance of panspermia makes the need of the "hardware" only second priority because, in such a case, Life is taken as granted. Obviously, rejecting panspermia, places problems of sites and conditions for the origin of life on Earth on the key position in projects leading to recognition of real origins of Life.

Many new proposal of chemical reaction or its sequences in prebiotic chemistry will certainly appear, fitting existing recognized sites of probable location of the reaction. One can hope also for new proposals concerning hardware, especially when at the time being the origin of life is still obscure.

Some references connected to the present summary, are listed in the subchapter in the book, by the present author (Zagorski 2010)

Acknowledgements

The membership in the Management Committee (2008–2012) of the European COST action CM0703 (Systems Chemistry) is acknowledged. The project is supported by the grant from the Polish Ministry of Science and Higher Education no. 365/N–COST/2008/0. (2008–2012).

References

Zagórski ZP (2007) Orig Life Evol Biosph 37:351–355

Zagórski ZP (2010) Chapter 5: role of radiation chemistry in the origin of Life, early evolution and in transportation through cosmic space. In: Basiuk VA (ed) Astrobiology: emergence, search and detection of life. American Scientific Publishers, USA, pp 97–154

E-mail: zagorski@ichtj.waw.pl

A Systems View to Define the Boundaries of the Problem of Origins of Life

Kepa Ruiz-Mirazo

Dept. Logic & Philosophy of Science—Biophysics Research Unit (CSIC-UPV/EHU) University of the Basque Country, Spain

Keywords System Properties • Minimal Biological Organization • (basic) Autonomy • Open-ended Evolution • Dynamic Decoupling

As this question suggests, we should indeed make a critical review of our ways of conceiving the "origin of life" problem, aiming to establish some common terms in which it could be posed or understood by the majority of researchers in the field. However, I disagree with the concrete proposal of doing so by referring, mainly, to the specific properties of the molecules involved (i.e., whether they are high molecular weight compounds or not). If we share the idea that life is not the property of a single molecule, but of a collection of complexly organized molecules, the boundaries of the problem of the origin of life should not be defined in terms of molecular specificities. Instead, we ought to think in terms of *system* properties, either general (if one considers life as a general property of many different systems) or specific (if one considers life as a special property of particular systems). In order to mark out the limits of the problem of origins of life, and assuming that such a complex transition cannot take place in a single-step event but it is a process that spans in time (over, let us say, millions of years), one should try to determine a point of beginning and a point of end of that process by characterizing the type of *organization* of the systems involved. In my contribution I will suggest two major properties or concepts to describe the organization of potential candidates for living systems, *autonomy* and *capacity for open-ended evolution*, which respectively mark out the beginning and the end of the process of origin of life. The first one is close to the idea of 'autopoiesis' (Maturana & Varela 1973), since it directly relates to the individual, metabolic (i.e., self-constructive) nature of life. The second is linked to the potential of a self-constructing system to construct other similar systems without an eventual decrease in its level of complexity (von Neumann 1966) and, although it has profound implications for the organization of individuals (phenotype-genotype decoupling), it involves a population of systems in collective evolution. My main claim (Ruiz-Mirazo et al. 2004; 2008) will be that the process of the origin of life begins with 'basic autonomous systems', possibly of low molecular complexity (i.e., made of low molecular weight compounds) but already organized as proto-metabolic cells; and it ends with 'genetically instructed metabolisms', or autonomous systems of high molecular complexity (high molecular weight compounds), also organized in cells but in a much more intricate, modular and hierarchical way: through a strong 'dynamic decoupling' made possible only with the development of a code of translation between genotype and phenotype. This decoupling is critical to combine the autonomy of individual living beings with their longer-term collective-evolutionary dynamics. So, according to such a scheme, open-ended evolution determines the conclusion of the process of origins of life and the aperture of the process of proper biological expansion and diversification.

References

Maturana H, Varela FJ (1973) De Máquinas y Seres Vivos—Una teoría sobre la organización biológica. Editorial Universitaria S.A., Santiago de Chile

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

Ruiz-Mirazo K, Umerez J, Moreno A (2008) Enabling conditions for open-ended evolution. Biol Philos 23(1):67–85

von Neumann J (1966) Theory of self-reproducing automata. In: Burks AW (ed) University of Illinois, Urbana

E-mail: kepa.ruiz-mirazo@ehu.es

OPEN QUESTIONS ON THE ORIGINS OF LIFE 2009