Viva Origino 34 (2006) 40-50

THE PROTEIN INTERACTION WORLD HYPOTHESIS OF THE ORIGINS OF LIFE

Peter Andras[1*] and Csaba Andras[2]

[1]School of Computing Science, University of Newcastle,

Newcastle upon Tyne, NE1 7RU, UK

[2]Department of Technology and Natural Sciences,

Sapientia – Hungarian University of Transylvania,

Miercurea Ciuc, Romania

(Received 16 December 2005 Accepted 26 December 2005)

(Abstract)

We present in this paper the protein interaction world hypothesis about the origins of life. According to this hypothesis life originated as a self-reproducing and expanding set of molecular interactions, and living cells are such sets of molecular interactions. We emphasize that according to our hypothesis the reproduction of molecular interactions is the critical aspect of such systems, while other origins of life hypotheses put the emphasis on the reproduction of molecules. We present molecular interaction systems in general, with examples presented in cellular context. We discuss the evolution of life forms according to our theory, and we show that RNA and DNA molecules can be seen as generators of molecular interactions that serve as memories of other such interactions in the context of cells viewed as molecular interaction systems.

Keywords: abstract communication system, molecular interaction system, origins of life, protein interaction, systems theory

1. Introduction

The question about the origins of life is still unanswered [1-3]. The most accepted theory proposes that life started in the form of RNA molecules (RNA – world) [2,4,5]. The most significant alternative theory suggests that life started with protein molecules (protein – world) [6,7]. Recently we described a further alternative the protein interaction world hypothesis (PIW – hypothesis) [8], according to which life may have started as an interaction system of proteins. The most significant difference between the protein – world hypothesis and the PIW – hypothesis being that the emphasis in the latter is on the reproduction of interactions, and not on the reproduction of molecules.

There is evidence suggesting that the synthesis of nucleic acids and sugars is possible in the prebiotic environment [4,9,10]. These molecules could have combined to form nucleobases and RNA molecules according to the RNA – world hypothesis [2,4], however the formation of nucleobases with sufficient yield in abiotic environments is questionable [11]. The ability of RNA molecules to catalyze chemical reactions may have led to self-replicating sets of RNA molecules, but again the low reliability of precise copying of RNA molecules may mean that such replicating sets of RNA molecules are not feasible [12].

Experimental evidence shows that proteins (peptides) can form in abiotic environments using protenoids [13] or thioesters [7]. According to the protein – world hypothesis such proteins may form self-reproducing sets of molecules [6,14-16]. The main problem of this hypothesis is that generally proteins lack the ability to replicate themselves (however note the replication ability of GADV-proteins [14]), so it remains unclear how protein-based life could store information about what to replicate. A promising variant of the protein – world hypothesis suggests that proteins and RNA might have co-evolved, the latter helping the reproduction of the former [6].

In this paper we present our alternative PIW – hypothesis [8], which emphasizes the importance of reproduction of molecular interactions. We revise to some extent our earlier view that suggested that life could have started with proteins only, and we suggest here that life could have started as a reproducing set of molecular interactions, which involved mostly protein and RNA molecules (moving closer to the position Lacey et al [6]). We apply the methodology of abstract communication systems theory to analyse the problem of emergence of life.

The paper is structured as follows: Section 2 describes molecular interaction systems according to the theory of abstract communication systems; in Section 3 we describe our interpretation of origins of life according to the PIW – hypothesis; in Section 4 we discuss implications of our theory; finally in Section 5 we close the paper with conclusions.

2. Molecular interaction systems

In this section we describe molecular interaction systems in the context of living cells. We follow the conceptualisation of abstract communication systems introduced in Andras and Andras [8].

Molecular interaction systems are considered as abstract communication systems. Such systems are built of communications generated by communication units. We consider molecules as communication units and their interactions as communications between these units. In our view the interaction system itself is made of these interactions, and excludes the molecules that generate these interactions as communication units. In cells, most molecular interactions involve proteins, and we view cells as mostly protein-based molecular interaction systems.

Molecular interactions depend on earlier interactions between molecules. A set of molecular

interactions that happened in a cell determine the range of further molecular interactions that may happen in the immediate or even later future. Similarly, in order for a molecular interaction to happen a set of earlier interactions should happen before. In other words, interactions between molecules generate a new set of molecules (possibly modified variants of the original molecules), and further possible interactions are those which are possible between the new molecules. Similarly, for an interaction between molecules to happen, the required molecules need to be present in the right conformation required for the interaction, which means that previous molecular interactions should have generated these required molecules. Such regularities that express the likelihood of interactions to happen following a set of earlier interactions are the continuation rules of molecular interaction systems. Regularities that express the fact that the realisation of an interaction depends on a set of necessary earlier interactions are the referencing rules of these systems. Referencing rules determine the set of interactions which are referenced by a current interaction, while continuation rules determine the range of interactions that may reference a set of earlier interactions.

A system of molecular interactions is a set of such interactions, which inter-reference frequently between themselves, and reference relatively rarely interactions outside of the system. Note that circular referencing of interactions is possible, as one interaction may be required for a second (i.e. it is referenced by the second one), and the second one may be followed by a reproduction of the first interaction (i.e. the second is referenced by the first). All other molecular interactions that are outside of the system (i.e. outside of the cluster of frequently inter-referencing interactions) constitute the environment of the system. The boundary of the system is defined in terms of the relatively sharp change in the frequency of inter-referencing between interactions. In case of cells molecules within the cell interact frequently with each other, generating an inter-referencing set of molecular interactions, which depend relatively rarely directly on molecular interactions that happen outside of the cell (e.g. execution of a signalling pathway needs the triggering possibly by an external molecule, which interacts with a receptor molecule, but most of molecular interactions involved in the pathway happen within the cell, and these interactions reference mostly other molecular interactions that happened within the cell). The boundary of the cell viewed as a molecular interaction system is materialized as the cell membrane or cell wall.

If a system of molecular interactions is not able to reproduce itself (i.e. to produce the set of interactions constituting the system again and again) it cannot be observed – most likely disappears before it is observed. Observable molecular interaction systems are those which reproduce. In addition those system which reproduce and expand (i.e. produce themselves in increasing quantity) are more likely to be observable than others that reproduce but do not expand. Systems compete with each other for molecular interactions in the sense of trying the integration of molecular interactions into the system, according to the referencing and continuation rules of the system. Systems which reproduce and expand faster than others may outcompete those which are slow in reproduction and expansion, and may drive the latter extinct. Cells are molecular interaction systems, which reproduce and expand. Those cells which are more successful in integration of molecular interactions into the system of the cell may drive to extinction other cells which are less successful in these terms (e.g. antibiotic resistant and non-resistant bacteria in the presence of antibiotics).

Memories in molecular interaction systems are such interactions that facilitate the reproduction of

themselves and also of a set of interactions that led to the original production of the memory interaction. For example, if the interactions A + B, C + D, E + F lead to the production of the interaction X + Y, and the interaction X + Y may lead to the reproduction of the interaction X + Y or to the reproduction of the set of interactions A + B, C + D, E + F then the interaction X + Y is a memory interaction in the context of the system which includes all these interactions. Memory interactions are critical for molecular interaction systems with high reproduction and expansion ability. Although in principle interaction systems may emerge without memory interactions and also based purely on memory interactions, the ability for reproduction and expansion of such systems would be very much inferior of systems, which involve both memory and non-memory interactions linked together by memories. In the context of cells we may consider the interactions of RNA molecules with themselves or other RNA molecules as memories of interactions between proteins (i.e. chains of amino acids). In particular, interactions of mRNA and tRNA molecules are memories of interactions between single amino acids and a chain of amino acids, such that the sequence of facilitated interactions leads to the production of a protein (the encoded by the mRNA). At the same time interactions of segments of DNA molecules may be viewed as memories of RNA interactions. For example the activation of a DNA segment is achieved by molecular interactions in which participate other DNA segments (promoter sites). The activation may lead to a sequence of molecular interactions which generate a premature mRNA molecule. During maturation this molecule interacts with other RNA molecules and generates further RNA molecules (e.g. siRNA, microRNA), which participate in further RNA interactions. In this way the interactions of the DNA segment facilitate the reproduction of interactions between RNA molecules.

Structures of molecular interaction systems are expressed additional constraints on interactions. Interactions conforming to constrained continuation and referencing rules follow the structures of the system. Such structures may materialize in context of cells as intra-cellular structural components (organelles), e.g. Golgi apparatus or ribosomes. Molecules participating in interactions that are part of structures are restricted in terms of their possible interactions with other molecules. Satisfying these additional constraints maintains the structures of the system. Having structures the system reduces the uncertainties of system rules, which in an environment where these structures are appropriate (i.e. it is possible to reproduce them sufficiently easily) implies increased ability of the system to reproduce and expand itself. In this way structures induce a simplification of the system (i.e. less variable rules) and at the same time contribute to the expansion of the system. Structures make the system specialized to fit a range of environment, increasing its reproduction and expansion ability in these environments. At the same a major change of the environment may render structures inappropriate, leading to a decrease in the system's ability to reproduce and expand. Structures also reflect constraints imposed by the environment onto the molecular interaction system. Environments with fewer constraints imply fewer emerging structures in systems.

Structures may lead to the emergence of inner subsystems of systems. Subsystems are clusters of molecular interactions, which satisfy the rules of the system, also satisfy a set of coherent structural constraints as well, and in addition constitute a set of frequently inter-referencing interactions surrounded by a referencing density boundary within the system. For example, molecular interactions of mitochondria

might be seen as a subsystem in the context of the cell's system. Mitochondrial interactions constitute a cluster of frequently inter-referencing interactions, which satisfy additional constraints to those of the cell, and surrounded by a boundary in terms of inter-referencing frequency. This boundary materializes in the form of the outer membrane of the mitochondria.

Molecular interaction systems may develop a special subsystem, which references memory interactions and generates new memories by processing other memories. This subsystem of the molecular interaction system is the information subsystem, which processes information about the system (i.e. memories of earlier interactions, which can be seen as information about the past of the system). Having an information subsystem increases the system's ability to reproduce and expand. This is because it increases the system's adaptation ability in sense of selection of appropriate mixes of system rules to be applied to generate / recruit new interactions in response to changes in the system's environment. In cell systems the subsystem of RNA and DNA interactions can be considered as the information subsystem. These interactions generate adaptively new RNA molecules (able to produce new memories), which engage in molecular interactions producing the right mix of proteins that is required for the reproduction and expansion of the cell in its changing environment. The interactions of RNA molecules and DNA segments constitute a subsystem of frequently inter-referencing interactions, surrounded by a density boundary in terms of frequency of inter-referencing. In more complex eukaryotic organisms (we consider complexity in terms of range of possible organismal behaviours) the core of the information subsystem constitutes a more segregated subsystem, the nucleus, the boundary of the subsystem being materialized as the nuclear membrane.

The information subsystem defines the identity of the system. The production of memories is controlled by the information subsystem, and memories determine (through facilitation) the molecular interactions that compose the system. In case of cells, their identity is defined by the interactions of their DNA and RNA molecules. Defining the identity of the cell also implies the definition of what is not considered to be part of the cell. Interactions that are not part of the cell's system are eliminated from the cell. For example, toxins or misfolded proteins may be neutralized or pumped out from the cell through interactions with other molecules.

Identity violations may take the form of faulty interactions (i.e. interactions which do not fit the interaction lexicon of the cell's system – e.g. interactions of a prion), error situations (i.e. when interactions obey the rule of the system, but their expected continuation interactions do not happen – e.g. halting growth because of lack of nutrients) or system failures (i.e. when the system experiences major shrinking – e.g. presence of toxins that cannot be eliminated form the cell). All these cases may indicate that the system does not match its environment sufficiently in order to reproduce and expand. In response to identity violations the molecular interaction system responds by adaptive change of its identity. Memories of the past of the system are processed and new memories are generated, redefining the identity of the system, with the aim of increasing the system's ability to reproduce and expand. In cells this means the adaptive change of the active parts of the DNA leading to changes in the range of active RNA molecules and proteins present in the cell. For example, bacteria may adapt to changing nutrient environments by switching on or off the production of proteins.

Molecular interaction systems compete with each other for molecular interactions that can be

incorporated into the system such that they satisfy the system's referencing and continuation rules (in other words, such that they fit the identity of the system). System interactions are essentially about defining what is the system and what is not (i.e. what should be incorporated and what should be expulsed). This also means that in a complementary sense they describe their environment (i.e. what is not the system). The ability of a system to reproduce and expand depends on how well it fits its environment. Systems with specialist subsystems are likely to outperform in terms of reproduction and expansion systems with less specialist structures. Systems having extensive information subsystem are also likely to outcompete systems with less developed information subsystems. At the same time being environmentally well adapted may also mean disadvantage in case of a major change in the environment. Specialist structures may become disadvantages in a changed environment. Highly developed information subsystem may help dealing with a wide range of environments, but at the same time may mean expensive overhead costs after a major environmental change, reducing the chances for reproduction and expansion of the system.

The complexity of molecular interaction systems can be evaluated in terms of their fit to their environment, which is in principle infinitely complex. Having structures, the system fits better its environments, implying that the simplification of the rules of the system imposed by the structures leads to a more complex system, which fits to a greater extent its environment than a less structured system. In the same way having an information subsystem means more elaborated identity and larger number of constraints on system-compatible molecular interactions. Again the simplification of rules leads to more complex systems that fit better their environment. Generally, more complex systems are likely to outcompete their less complex competitors in relatively stable environments. At the same time less complex systems are likely to suffer less in the case of major environmental changes, and following such events may gain advantage in the competition against more complex systems adapted to the wrong (the previous) environment.

3. The protein interaction world hypothesis

We consider living cells as systems of molecular interactions. In this sense, life started as a self-reproducing and quantitatively expanding molecular interaction system. The competition of such systems led to molecular interaction systems of extant cells.

To answer the question about the origins of life we need to find relatively simple molecular interaction systems that have the ability of self-reproduction and expansion. We note that Kauffman [17] and Segre and Lancet [3] proposed to some extent similar lines of reasoning about searching for the origins of life in terms of molecular interaction systems. Our discussion of protein interaction systems indicates that in principle it may be possible to build self-reproducing and expanding molecular interaction systems without having memories or from molecular interactions that all act also as memories. However in both cases the expected ability of the system to reproduce and expand would be very moderate. This suggests that molecular interaction systems able to start early life could have emerged as systems of molecular interactions supported by a core of memory interactions.

According to the protein interaction world (PIW) hypothesis life started as a molecular interaction system involving mostly interactions of proteins (peptides) with other proteins and possibly other molecules, and supported by molecular interactions of RNA molecules, which served as memory interactions for molecular interactions of proteins. It is important to emphasize that living cells are seen as systems of molecular interactions. The actual

molecules are not considered as part of the system, but only as communication units, which generate interactions that constitute the system. This means that the reproduction of all molecules within the system is not required according to our hypothesis, and only the reproduction of interactions between molecules is required. In principle the molecules may be produced outside of the narrower context of the system. Only interactions of molecules are incorporated into the system according to the rules of the system, when these molecules are available to produce these interactions.

We believe that life originated in hydrothermal volcanic marine environments. Experimental evidence shows that amino acids can form tight clusters called proteinoids at high temperatures [6], which may lead to the formation of peptides [18]. An alternative way of building peptides is by the transformation of thioesters [7,19], a chemical pathway that works efficiently in abiotic conditions. Experimental simulations of marine hydrothermal vents have shown that amino acids may form short peptides in such conditions [15]. Recently, Leman et al [13] have shown that peptides may form with high yield in volcanic marine environments in the presence of carbonyl sulphide, a common volcanic gas. Genetic analysis evidence also suggest that early life emerged in high temperature environment rich in sulphur [20], which implies the plausibility of the above mentioned ways to the synthesis of early peptides. Interactions between peptides may catalyse the synthesis of fatty acids, lipids, sugars and nucleic acids, some of which are expected to be present in early abiotic environments [21]. Pores and micro-tunnels in the submarine surface may have created the appropriately dry micro-environments and sufficiently high concentrations of organic molecules such that self-reproducing molecular interaction systems could have emerged.

In peptide-rich environments of volcanic hydrothermal vents proto-cells could have emerged as molecular interaction systems, which rely on the production of required proteins (communication units) elsewhere in their larger environment, and reproduce themselves by recruiting molecular interactions that fit their regularities. In case of such proto-cells the boundary of the system could be enforced by their environment in form of physical separation of pores and micro-tunnels. Out of the possible early forms of self-reproducing protein interaction systems those could become dominant and fastest expanding, which relied on the combination of protein interactions and memories of such interactions in form of interactions of RNA-like molecules.

Among proto-cells those are most likely to expand, which can expand by recreating their boundary easily. Reproducing fatty membranes forming a lining of walls of pores and micro-tunnels might have led to the production of membranes of proto-cells, making them able to recreate and enforce their own system boundary. The emergence of such systems of molecular interactions could constitute the first step towards modern cells.

Our discussion of molecular interaction system shows that those systems are likely to succeed in more constrained environment, which develop appropriate structures and subsystems. This means that as early life-forms conquered more hostile environmental niches, outside of the supportive environments of hydrothermal vents, environmental constraints induced structural constraints on proto-cells forcing them to develop structures and related subsystems. The same environmental constraints and variations could have also induced the emergence of the information subsystem in proto-cells, which materialized in the form of increasing number of RNA interactions supported by the emergence of the incorporation of DNA molecule interactions as

memories of RNA interactions. Cells with structures, subsystems and in particular with information subsystem could have easily driven to extinction possible earlier more primitive forms of life.

According to our hypothesis more complex life forms developed by applying further simplifications to their rules by further structures and growing information subsystem. These cells gained in complexity and reproduction and expansion ability in their environment. This leads to the emergence of eukaryotic cells with well separated nucleus enclosing the core of their information subsystem. Eukaryotic cells are able to build complex multi-cellular organisms increasing further their ability to reproduce and expand. In a similar manner, an increasing regulatory component of the DNA (e.g. reaching up to 95% of the total DNA in some complex organisms) led to further increase in environmental fitness and complexity of these organisms.

Summarizing our hypothesis, life emerged as a self-reproducing and expanding set of molecular interactions in hydrothermal volcanic marine environments. The key players of these interactions were proteins (peptides) and RNA-like molecules, which generated interactions representing memories of protein interactions. In our view the critical aspect of living system is the reproduction of molecular interactions and not necessarily the local reproduction of molecules. Starting from pores of rocks enclosing reproducing sets of interactions, life evolved by building independently lipid membranes, creating early forms of cellular membranes. Conquering less friendly environments led to the emergence of structures and expansion of the information subsystem of proto-cells. Cells with cellular organelles (structures) and DNA-based RNA interaction system emerged as dominant life forms. Such cells developed further by applying further simplifications to their rules through structures and developing information subsystem and gaining in complexity and ability to reproduce and expand in their environment.

4. Discussion

4.1. Evolution of the amino acid dictionary

It has been discovered recently [22] that amino acid replacement likelihoods are not symmetrical during phylogenetic evolution. The data shows that sulphur containing amino acids are increasing in their frequency, while ‘old’ amino acids (i.e. those which could have been present in early abiotic environments) are decreasing in their frequency, which is in agreement with the commonly accepted view is that life started with a few amino acids and newer ones were added to the amino acid library of organisms during evolution [22].

Considering that there are in total 22 amino acids used in living organisms [23] and that 2 out of these are rarely used (seleno-cystein and pyrolysine), in our view, the asymmetric replacement likelihoods of amino acids indicate a simplification of the amino acid dictionary that is still going on. Considering that plants and bacteria produce a wide range of amino acids which are not used as genetically encoded components of proteins [24], it is possible that at the beginning of life there was a much wider lexicon of amino acids that was allowed to be used as memorised components of early proteins. The possibility of encoding more than 22 amino acids by the usual triplets of nucleobases constituting codons, the possibility of using four nucleobase codons [25], and the fact that the smallest tRNA sets of mitochondria contain 22 tRNAs, while larger sets might contain more than 60 tRNAs, all suggest that in principle it could have been possible the use of a larger amino acid lexicon by earlier life forms. Our systems theoretical discussion of molecular interaction systems indicates that such systems could

have gained in their ability to reproduce and expand by simplifications of their rules, which in turn led to more complex behaviour and better match to their environment. This suggests that it is possible that indeed the asymmetric replacement likelihoods reflect an ongoing process of simplification of the amino acid lexicon.

Our theory also suggests an explanation of the fact that sulphur containing amino acids are gaining in terms of frequency through the asymmetric replacements. In our view such changes should reflect a better fit to the environment. Considering the sulphur rich environment of early life forms it is likely that in order to increase their fitness to their environment they needed to increase the frequency of sulphur containing amino acids in their proteins. Possibly this early process of preference for sulphur containing amino acids continues today in the cells of extant life forms.

A way to test the validity of our interpretation would be to analyse amino acid replacement likelihoods in the context of thermophile unicellular organisms, which live in a particularly sulphur rich environment. Our expectation is that in these organisms we should see an even more imbalanced replacement likelihood assignment, which would favour even more the replacement of non-sulphur containing amino acids with sulphur containing amino acids.

4.2. RNA interactions

Our hypothesis about the origins of life suggests that in order for the information subsystem to develop an increase in the amount of RNA interactions was necessary. Such interactions would constitute memory processing interactions, generating new RNA molecules able to facilitate new protein interactions through their interactions with other molecules.

The relatively recent finding that non-coding RNA molecules are plentiful [26], and that microRNAs and siRNAs play an important role in regulating the maturation and translation of mRNA molecules [27], indicate that indeed there is a major role played by RNA interactions in living cells. Other recent results [28] show that it is possible to restore mutated DNA by relying of RNAs and their interactions, indicating again the importance of RNA interactions and also the link between RNA interactions and their DNA memories.

According to our theory these results fit the expectation that underlying the system of protein interactions there is extensive system of RNA interactions. Our theory also suggests that the system of RNA interactions should be more extensive in more complex organisms, and should be less developed in more primitive organisms. The theory also predicts that more extensive RNA interaction system should correlate with the presence of larger amount of regulatory DNA segments. Furthermore, considering that RNA interactions express constraints on protein interactions, we expect that primary response to changing environmental conditions should happen in terms of changing patterns of RNA interactions, which leads to adaptive changes in terms of protein interactions.

5. Conclusions and remarks

We presented in this paper the protein interaction world (PIW) hypothesis of the origins of life. According to this hypothesis life emerged as a set of molecular interactions able to reproduce and quantitatively expand themselves. Our hypothesis is similar to some extent to variants of the protein – world hypothesis; however the main emphasis in our hypothesis is on the reproduction of molecular interactions and not on the reproduction of molecules.

We discussed how life might have emerged and developed according to the PIW hypothesis. We

pointed out that the role of RNA molecules according to our hypothesis is to produce molecular interactions that act as memories of protein interactions in the sense of facilitating the reproduction of such interactions. In a similar manner, we suggested that DNA molecules generate molecular interactions that are memories of RNA interactions. We described the system of RNA and DNA interactions as the information subsystem of the cell considered as a molecular interaction system.

We did not discuss the energetic feasibility of our hypothesis. We considered that the conceptual simplicity of our hypothesis (i.e. everything is formulated in terms of sets of inter-referencing interactions, memories and constraints on referencing and continuation rules) makes our hypothesis interesting and appealing. We are aware that energetic considerations imply significant practical restrictions on origins-of-life hypotheses, but we are also aware that unusual or extreme conditions may help in overcoming otherwise impassable energetic barriers. Such special conditions might have been present in the environment of early life (e.g. high temperature, high sulphur content, semi-dry environment). The energetic feasibility analysis of our theory is part of our future work plans.

A critical aspect of any hypothesis is that it should be able to lead to testable predictions, which can be checked leading to validation or invalidation of the hypothesis. In case of the PIW hypothesis such potentially testable predictions are: (1) organisms living in high sulphur content environment should have an increased frequency of sulphur containing nucleobases and amino acids in their RNAs and proteins, respectively, than the average frequency of such nucleobases and proteins in organisms; (2) the range and amount of RNA interactions should be higher in more complex organisms than in less complex organisms (where complexity is measured in terms of range of distinct behaviours); (3) similar behavioural complexity of two species, such that one has more developed RNA interaction system, should imply that this one has a more constrained set of protein interactions (i.e. possibly fewer expressed proteins or fewer basic types of interactions, which are organised into more complex spatio-temporal patterns of interactions).

References

1. Ganti, T. Biogenesis itself. *Journal of Theoretical Biology*, **187**, 583-593 (1997).
2. Joyce, G.F. Booting up life, *Nature*, **420**, 278-279 (2002).
3. Segre D. and Lancet, D. Composing life, *EMBO Reports*, **1**, 217-222 (2000).
4. Joyce G.F. RNA evolution and the origins of life, *Nature*, **338**, 217-224 (1989).
5. Zubay, G. and Schechter, A. Current status of the RNA – only world, *CHEMTRACTS – Biochemistry and Molecular Biology*, **13**, 829-836 (2000).
6. Lacey, J.C., Cook, G.W., and Mullins, D.W. Concepts related to the origin of coded protein synthesis, *CHEMTRACTS – Biochemistry and Molecular Biology*, **12**, 398-418 (1999).
7. DeDuve C. RNA without protein or protein without RNA ? In: *What is Life ? The Next Fifty Years*, Murphy, MP, O'Neill, LAJ (eds.), Cambridge University Press, Cambridge, UK, pp.79-82 (1993).
8. Andras, P. and Andras C.D. Protein interaction world – an alternative hypothesis about the origins of life, *Medical Hypotheses*, **64**, 678-688 (2005).
9. Glavin, D.P. and Bada, J.L. Isolation of purines and pyrimidines from the Murchison meteorite using sublimation, *35th Lunar and Planetary Science Conference*, paper 1022 (2004).

10. Robertson, M.P. and Miller, S.L. An efficient prebiotic synthesis of cytosine and uracil, *Nature*, **375**, 772-774 (1995).
11. Dworkin, J.P., Lazcano A., and Miller, S.L. The roads to and from the RNA world, *Journal of Theoretical Biology*, **222**, 127-134 (2003).
12. Maynard-Smith, J. and Szathmary, E. *The Major Transitions in Evolution*, Oxford, Oxford University Press (1997).
13. Leman, L., Orgel, L., and Ghadiri, M.R. Carbonyl sulphide mediated prebiotic formation of peptides, *Science*, **306**, 283-286 (2004).
14. Ikehara, K. Possible steps to the emergence of life: the [GADV]-protein world hypothesis, *The Chemical Record*, **5**, 107-118 (2005).
15. Imai, E., Honda, H., Hatori, K., Brack, A., and Matsuno, K. Elongation of oligopeptides in a simulated submarine hydrothermal system, *Science*, **283**, 831-833 (1999).
16. Kawamura K., Kuranoue K., and Nagahama M. Prebiotic inhibitory activity of protein-like molecules to the template-directed formation of oligoguanylate from guanosine 5 '-monophosphate 2-methylimidazolide on a polycytidylic acid template, *Bulletin of the Chemical Society of Japan*, **77**, 1367-1375 (2004).
17. Kauffman, S.A. 'What is life ?': was Schrodinger right ? In: *What is Life ? The Next Fifty Years*, Murphy, MP, O'Neill, LAJ (eds.), Cambridge University Press, Cambridge, UK, pp.83-114 (1993).
18. Francis, B.R. A hypothesis that ribosomal protein synthesis evolved from couple protein and nucleic acid synthesis, *CHEMTRACTS – Biochemistry and Molecular Biology*, **13**, 153-191 (2000).
19. Deamer, D.W. and Chakrabarti, A. The first living organisms: In the light or in the dark, *CHEMTRACTS – Biochemistry and Molecular Biology*, **12**, 453-467 (1999).
20. Koonin, E.V. Comparative genomics, minimal gene sets and the last universal common ancestor, *Nature Reviews Microbiology*, **1**, 127-136 (2003).
21. Miller, S.L. and Orgel, L.E. *The Origins of Life on the Earth*. Englewood Cliffs, NJ, Prentice Hall (1974).
22. Jordan, I.K., Kondrashov, F.A., Adzhubei, I.A., Wolf, Y.I., Koonin, E.V., Kondrashov, A.S., and Sunayev, S. A universal trend of amino acid gain and loss in protein evolution, *Nature*, **433**, 633-638 (2005).
23. Cobucci-Ponzano, B., Rossi, M., and Moracci, M. Recoding in Archaea, *Molecular Microbiology*, **55**, 339-348 (2005).
24. Singh, B.K. (ed) (1999). *Plant Amino Acids.* New York, Marcel Dekker (1999).
25. Magliery, T.J., Anderson, J.C., and Schultz, P.G. Expanding the genetic code: Selection of efficient suppressors of four-base codons and identification of "shifty" four-base codons with a library approach in Escherichia coli, *Journal of Molecular Biology*, **307**, 755-769 (2001).
26. Barciszewski J. and Erdmann V.A. *Noncoding RNAs : molecular biology and molecular medicine*, New York, Kluwer Academic / Plenum Publishers (2003).
27. Denli, A.M., Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., and Hannon, G.J. Processing of primary microRNAs by the Microprocessor complex, *Nature*, **432**, 231-234 (2004).
28. Lolle, S.J., Victor, J.L., Young, J.M., and Pruitt, R.E. Genome-wide non-mendelian inheritance of extra-genomic information in Arabidopsis, *Nature*, **434**, 505-509 (2005).