Life without the double helix: Oparin Revisited

by Doron Lancet
Department of Molecular Genetics and the Crown Human Genome Center, The Weizmann Institute of Science, Rehovot, Israel

Viewing a Watson-Crick model of DNA is an exhilarating experience. It becomes crystal-clear how the base-paired double helix allows information to be stored and copied. Yet, in the post-1953 euphoria, scientists were tempted to believe that having such an explicit chemical entity as a functional core of present-day life implies that life began with it. The RNA-world model conjectures that the first chemical entity capable of self-replication was a base-paired polynucleotide. The present paper highlights reasons why this is unlikely to be right, and delineates an alternative scenario that draws upon Oparin's early teachings.

From structure to origin

A living cell is an utterly complex molecular machine. Origin of life scientists hope that by understanding its intricacies they will stumble upon an explanation for how it emerged. This aspiration is based on experience with simpler systems. For example, by studying moon rocks it is possible to come up with a mechanism of its formation by a cataclysmic planetary collision. However, this may not be true for entities whose complexity is many orders of magnitude higher. In such cases it is necessary to employ divergent thinking, guided by an assessment of which features are fundamental and which are consequential.

It is widely agreed that the most crucial property of living entities is a capacity to generate similes. This attribute underlies the process of evolution and natural selection, and must have been present very early in life's history. An often stated chicken-and-egg conundrum is that life cannot reach even a most rudimentary level of complexity without self-replication, but that only minimally

An often stated chicken-and-egg conundrum is that life cannot reach even a most rudimentary level of complexity without self-replication, but that only minimally complex chemical entities can multiply.

Alexander Ivanovich Oparin (1894-1980), Russian biochemist. Oparin graduated from Moscow University in 1917 and became professor of plant biochemistry there in 1929. He also helped to found the Bakh Institute of Biochemistry. As early as 1922 Oparin was speculating on the origins of life on Earth and later published The Origin of Life on Earth (1936) and The Chemical Origin of Life (1964).

Science Photo Library

complex chemical systems can multiply. The discovery of DNA, simple in design and seemingly capable of creating its own copies, has led to a wide belief that nucleic acids provide a solution for such an incongruity.

However, an examination of the structure of DNA and RNA reveals chemical intricacy that defies the notion that they were the first replicators. Even in the present-day setting, polynucleotides only undergo replication in the context of its cellular milieu. In a primordial chemically-heterogeneous scenario, the spontaneous emergence of sufficient quantities of the required carbohydrate moiety and nitrogen base components is unlikely. Also, phosphodiester polymerization is thermodynamically unfavorable, and the repeated building of a long second nucleic acid strand on its first strand template would require intricate catalysis [1, 2]. For these and other reasons it has been argued by quite a few authors that an RNA-like polymer may not have been the early seed of life [3, 4].

The breakthrough discovery of Watson and Crick [5, 6] led to the practical extinction of alternative hypotheses. The beauty of DNA replication was so compelling that any other "dirtier" pathway suddenly looked untenable. Yet, in view of the problems posed above, it is interesting to ponder the untainted state of art of origin-of-life research as it existed prior to 1953. In the process, insight may be gained with respect to the dichotomy between the widely accepted RNA-world model and an emerging, more debatable lipid world model for life's emergence *(Fig. 1)*.

Oparin revisited

The Russian biochemist Aleksandr Ivanovich Oparin first published his pamphlet *The Origin of Life* in 1924 [7](reprinted in [8]). It was only a later expanded version published as a book [9] that got translated to English shortly thereafter [10, 11]. In that book, Oparin dealt with the problem of life's origin, for the first time, from a materialistic perspective not inhibited by the religious constraints prevailing in Western nations. The standard Oparin-Haldane theory [12] puts forward the production of organic molecules on the early Earth followed by chemical reactions that produced increased organic complexity. This process is perceived to have led eventually to organic life capable of reproduction, mutation, and selection. In this form, the theory has no specification of which molecules were first, or how self replication came about. In fact, it predates by two decades the Watson-Crick breakthrough, and is therefore unbiased by its implications.

The origin of organic molecules and the origin of replication are two separate problems. Actually, the first may be viewed as a prerequisite for the second. The groundbreaking experiments of Miller and Urey [13], first published in the same year as DNA's structure, were specifically aimed at the first question. They asked how, under primitive earth conditions, a sufficient supply of organics was generated to warrant Oparin's "primordial soup", the milieu in which reactions leading to life are assumed to have occurred. Criticism has often been aimed at both Miller and Oparin for conjecturing a particular set of atmospheric conditions for the generation of an organic soup, conditions that may have not actually prevailed historically. The present paper avoids this controversy, positing that organic compounds could have arisen by a wide variety of processes under a broad spectrum of planetary conditions [14-23]. We address only the ways by which the double helix's discovery affected scientific views on the earliest replicators.

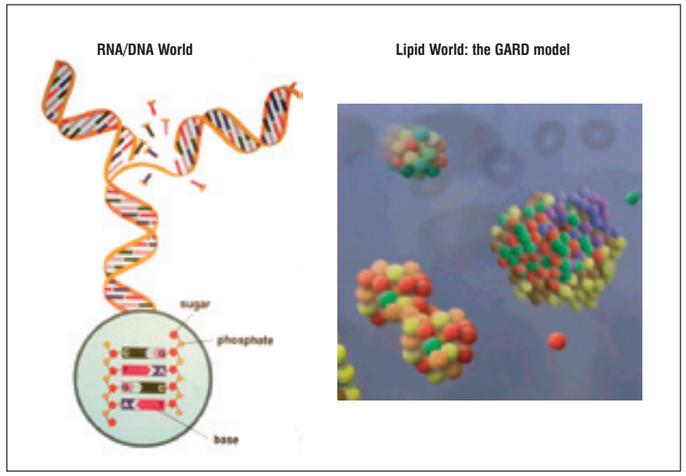
According to Oparin's teachings the following steps have happened en route to early life:

1) Random synthesis of simple organic molecules from atmospheric gases.

2) Formation of larger, more complex molecules from the simple organic molecules.

3) Formation of coacervates - unique droplets containing the different organic molecules.

4) Development of a capacity to take up and discharge specific molecules, and maintain a characteristic chemical pattern or composition.



5) Development of "organizers" that allowed controlled reproduction to ensure that resultant split-related daughter cells have the same chemical capabilities.

6) Beginnings of evolutionary developments so that a group of cells could adapt to changes in the environment over time.

This clear-cut hierarchical scenario was considered for decades an acceptable view of how life began. In later years [24] Oparin had guessed that the hypothesized "organizers" in step 5 might include nucleic acids, but clearly he was not thinking in terms of a full-fledge transcription and ribosome-mediated translation apparatus, as these would have been too complex for the primitive stages invoked. Oparin had the insight to consider the possibility of simile generation based on a web of chemical interactions, namely what would be called nowadays a prebiotic metabolism-like network [25-27]. Self replication was thought of as related to the splitting of coacervate droplets, including its entire molecular repertoire, a process conceptually similar to modern cell division. Oparin was not influenced by the later realization that within each cell lurks an individual component, DNA, capable of information storage and molecular self-replication.

The nucleic acid takeover

In the post-1953 era, origin of life research underwent a revolution, not less significant than that which affected all of molecular biology. This revolution hinged on the notion that base-paired nucleic acid polymers might be capable of generating their own copies. Thus, the vaguer, but potentially more

Figure 1

The two dichotomous stereotyped "world" scenarios for Life's origin. Both are likely to be over-simplifications, since prebiotic mixtures must have included diverse compounds, and not just one class of substances. Therefore, the two scenarios should be taken to represent opposing concepts. On the one hand, the "RNA/DNA world" view asserts that life began with individual molecules that had a relatively complex covalent structure, enabling them to store information, instruct the synthesis of other molecules, and undergo a complex chain of chemical steps that resulted in their own replication. On the other hand – the "lipid world", a scenario that claims that much before the appearance of RNA-like biopolymers, there existed "non-covalent polymers"[40], assemblies of relatively small molecules, capable of mutual association by weak chemical forces, including, but not limited to hydrophobic interactions [34]. In the realm of this scenario, information was contained in an assembly's chemical composition rather than in a molecular covalent sequence (Fig. 2), and replication was based on an assembly's random fission rather than on orderly templating. In the minds of lipid world proponents, never was there an early bottleneck in which one individual molecule was responsible single-handedly for practically all pivotal life-promoting functions. Since it is straightforward to calculate that uncontrolled organic chemistry may easily produce 10^7 different compounds, and since the simplest free-living cell has about 1,000 molecule types, the RNA world may be typified by a molecular count time-sequence of 10^7, 1, 10, 100, 1000 while the lipid world would have a sequence 10^7, 10^6, 10^5, 10000, 1000. Future research would determine which path is more realistic.

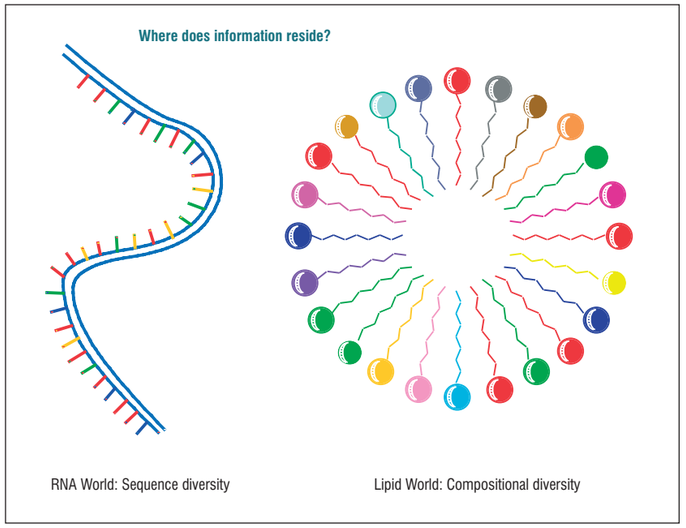


Figure 2

In covalent biopolymers, information resides in the sequence of chemical units (monomers), while in non-covalent molecular assemblies it rests in a "compositional vector" whose elements are the counts of the different molecular types. See refs. 35, 37, 40.

realistic scenario of a "mixed bag" ("garbage bag" in Freeman Dyson's terminology [28]) that undergoes self-reproduction has been largely abandoned.

The new scenario was simple and seemingly elegant. If only a single self-replicating base-paired nucleic acid could form in the soup, life would emerge. This is based on the notion that once self replication is in place, selection and evolution would be jump-started and then gradual improvements could lead from a naked nucleic acid to a protocell. This scenario envisions that the additional paraphernalia needed for a full-fledged cell-like entity, such as metabolic pathways for the synthesis of hundreds of molecular components, membrane enclosures, passive and active transporters, energy-harvesting mechanisms and ATP-like free energy-rich compounds, primitive tRNA's and their associated sythetases, ribosomes and protein enzymes would all somehow appear and join the original nucleic acid. Typically, it is not specified how the fact that a self-replicating nucleic acid has already turn up aids the appearance of the other molecular constituents. Also, little is said about how, if a single molecular replicator is about, it will recruit all the other components into the replication, selection and evolution game.

The polynucleotide-first scenario brings to mind Marcia Brown's popular children story, based on an old French tale, about three hungry soldiers who manage to cook up a fine soup by cajoling the necessary ingredients from a village of wary French peasants. The pot initially contained only a stone immersed in water, and of course the soup emerged irrespective of it, solely because meat, vegetables, salt, spices and all the rest have been stealthily added. The early evolution moral is that the "recruitment by nucleic acids" picture may be somewhat simplistic, and that alternative, Oparin-style, scenarios need to be explored. In the latter, the hypothesized early entities would include a diverse repertoire of components right from the start, and full-fledged double-helix-based mechanisms would arise much later. Replication could emerge as a capacity of an entire molecular ensemble, so that evolution by gradual improvement would apply to all components in unison, and not solely to a hypothesized core information carrier.

Templates and catalysts

One of the most popular means of portraying the chicken and egg nature of early evolution is asking what came first: nucleic acid-based information storage molecules needed in order to synthesize proteins, or globular protein catalysts without which the DNA/RNA machinery would come to a grinding halt. One way to answer the question is to ask which of the two sets of chemical phenomena appears more basic. The RNA world protagonists' answer is that since ribonucleic acids are capable of both template-based replication and catalysis, then RNA must have been the primordial mover. However, a view derived from basic chemistry may lead to a different answer. This alternative view probes the very basic definition of a templating reaction. As seen in *Fig. 2*, all a templating strand is doing when directing the incorporation of a properly base-paired nucleotide into a growing second stand is a special case of rate enhancement. While this reaction does not conform to the strict definition of multiple turn-over enzyme reaction, it is nevertheless a selective catalysis phenomenon. One may envisage a very primitive system in which the prefect Watson-Crick structure has not yet evolved, and in which the control over a replication-related reaction is in the hands of a more unwieldy set of organic catalysts. Addressing the chemical nature of templating in such a general perspective leads to the notion that sets of mutually catalytic substances may be, in principle, sufficient for primitive forms of self replication.

It is possible to envisage how through a long, gradual process of selection and evolution the double helix would emerge from its "bag of catalysts" ancestor. Another way to view this is to say that a fairly complex network resembling present day metabolism is more basic, therefore likely to have been earlier than a double helix-based genetic apparatus. Indeed, Oparin's original paper [7, 8] makes the argument that "slowly but surely, from generation to generation, over many thousands of years, there took place an improvement…directed towards increasing the efficiency of the apparatus for absorption and assimilation of nutrient compounds… (and) the ability to metabolize".

Primordial Information storage

Oparin's first article was written before Avery's discovery of DNA's role as a genetic information store [29, 30], and much before Watson and Crick uncovered the mechanism by which this is made possible. To nucleic acid supporters, Oparin's description of how biological information was stockpiled and transmitted looks rather primitive. He describes how bits of protoplasm-like 'gel' would be "…broken down and give rise to new pieces…constructed or organized just like their parent…and therefore their structure was inherited by them from the gel from which they were formed". The essence of this argument is that the most important aspect of replication is just bequeathing the overall structure and function of a molecular ensemble through a capacity to propagate an internal chemical consistency, plus a simple process of physical split.

In fact, a purist's point of view would assert that this is true also for a present day living cell. Accordingly, the molecular templating of DNA is just one of a large set of mechanisms through which it is assured that after a physical split the two cellular halves would be capable of replenishing their molecular repertoires en route to the next split. Parallel mechanisms would be protein synthesis, the assembly ribosomes, the protein-mediate copying of the centrioles

Dr Gopal Murti/ Science Photo Library

DNA replication.

Coloured Transmission Electron Micrograph (TEM) of human DNA from a HeLa cancer cell, showing a stage of DNA replication. The strand of DNA is coloured yellow. It has formed into a Y-shaped molecule termed a replication fork, where the DNA has unwound into two single strands. Normally, DNA consists of two tightly wound spiral strands. During replication, a "bubble" region forms which enlarges to form a replication fork. It is here that daughter strands form as the parent DNA acts as a template for the construction of a new matching strand. In this way the sequence of bases (or genetic information) along the DNA molecule is replicated.

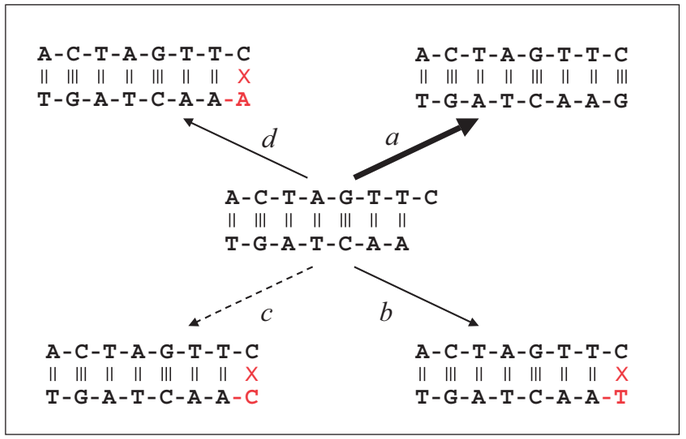


Figure 3

It is proposed that templating is derived concept, and the more basic notion is catalysis. The elementary step of adding a base to an extending second strand in a double stranded RNA/DNA type biopolymer may be summarized in four competing reactions, one which involves legitimate Watson-crick pairing (*a*), and the other three (*b,c,d*) forming illegitimate pairing with decreasing kinetic efficacy. The top strand may be considered to be an "enzyme" that catalyses chemical reaction *a* in favor of the side reaction reactions *b, c* and *d*.

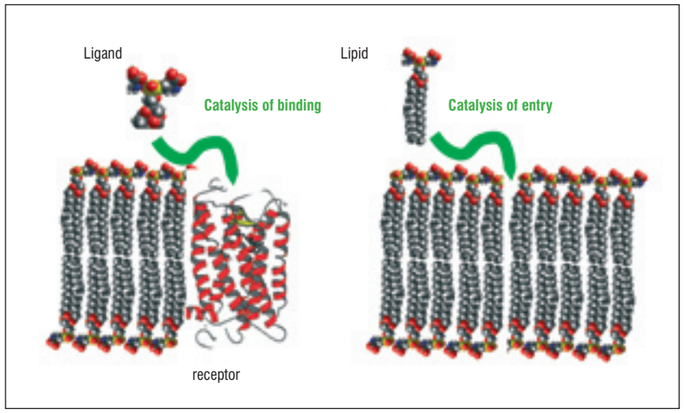


Figure 4

Left, Membrane lipid phase as catalyst for peptide-receptor interactions [41]. Right, in analogy, neighboring lipids in a micelle or a bilayer may catalyse the entry/exit of another molecule into the amphiphilic assembly. Such a relatively simple notion serves as a core of the GARD model as described in the text and elsewhere [33, 40].

and the production of new lipid bilayers for organelles such as mitochondria and the Golgi apparatus.

While the words of Oparin describe vividly how structure and function might have been passed down the generations, a more accurate and rigorous description would be needed to convince a modern molecular biologist that the split of a molecular assembly is sufficient to do the information transfer job over many generations. To this end, we have advanced in recent years the Graded Autocatalysis Replication Domain (GARD) model [31-35]. The essence of this model relates to the capacity of a daughter molecular assembly, derived by a physical split from its ancestor, to withstand the "trauma" of fission (e.g. the loss of scarce but essential molecular species). GARD molecular assemblies are shown to be capable of undergoing a complex replenishment process during growth, thus assuring homeostatic growth and a capacity for self replication upon fission [36, 37].

In a modern-day cell, this process is known to depend crucially on protein synthesis, based on information inscribed in RNA and DNA. Such central dogma pathway is, however, so complex that it is unlikely to have been present in very early protocells [25, 38, 39]. The RNA world concept asserts the existence of very few, template-replication biopolymers, assisted by an ill-specified set of catalysts, potentially RNA itself, to make proteins. The GARD scenario claims that early on, none of the three components of the central dogma may have existed, and that the only resemblance of early replicating and evolving entities to present day cells was in matters of general design.

Specifically, GARD assumes a network of mutually catalytic, relatively simple organic molecules, with amphiphilic characteristics, forming small assemblies. The most rudimentary form of rate enhancement is assumed to have occurred upon amphiphile entry into and exit out of the molecular assemblies *(Fig. 3)*, resulting in a behavior that resembles cellular homeostatic growth [33, 37, 40]. More advanced versions of GARD currently being developed, include also catalyzed covalent bond formation that leads to oligomerization and to the emergence of a more realistic connected metabolism. In this view, simple catalysts (enzyme mimetics) underlie the emergence of primitive catalytic networks, and the eventual appearance of DNA simply subserves the making of better

catalysts, folded polypeptides necessary to generate highly effective enzymes.

Computer simulations, that embody a set of rigorously-defined kinetic and thermodynamic constraints, demonstrate the capacity of such assemblies to grow homeostatically, i.e. while preserving their compositional information *(Fig. 4)* [37], a prerequisite for Oparin-style recovery from fission. Thus, GARD assemblies appear to alleviate the chicken-and-egg problem, harboring within them both replicable information and a metabolism-like network.

Conclusions

As we proceed into the post-genomic era emphasis is being shifted from DNA sequences to the more global molecular interplay in the realm of systems biology. Similarly, it is hard to envision primordial cellular systems without considering some form of mutually interacting sets of compounds. It is possible, however, that delineating the progression from a simple protocell to the "last universal common ancestor" of terrestrial life, including double-helical DNA, would be as convoluted as outlining how human beings evolved from a lowly worm.

The 1953 ground-breaking discovery of DNA structure brought with it tremendous progress with respect to our understanding how present life is molecularly constructed. Paradoxically, it may have changes the course of prebiotic evolution research in a somewhat less favorable fashion. Time will tell whether proponents of the double helix-first scenarios are right or wrong. But surely, the unbelievable aesthetic values of DNA's structure made people focus on one origin of life scenario more than truly warranted. A broader vista, in which room is made for diverse molecular mechanisms, while still paying tribute to the all-important template-based replication, might be beneficial for the ultimate cracking of what may be the most important open question in "low energy" science – how life emerged from inanimate chemical compounds. ■

Acknowledgements

I wish to thank Ran Kafri for contributing ideas, Barak Shenhav for critically reading the manuscript and Daniel Segré for groundbreaking research and help with the illustrations.

References

1. Bartel, D.P. and P.J. Unrau, Constructing an RNA world. Trends in Biochemical Sciences, 1999. 24(12): p. M9-M13.

2. Ertem, G. and J.P. Ferris, Synthesis of RNA oligomers on heterogeneous templates. Nature, 1996. 379(6562): p. 238-40.

3. Schwartz, A.W., The RNA World and its origins. Planetary and Space Science, 1995. 43(1-2): p. 161-165.

4. Shapiro, R., The improbability of prebiotic nucleic acid synthesis. Orig Life, 1984. 14(1-4): p. 565-70.

5. Watson, J.D. and F.H.C. Crick, Genetical implications of the structure of deoxyribonucleic acid. Nature, 1953. 171(964-967).

6. Watson, J.D. and F.H.C. Crick, A structure for deoxyribose nucleic acid. Nature, 1953. 171: p. 737-738.

7. Oparin, A.I., Proiskhozhdenie zhizny. 1924, Moscow.: Izd. Moskovski Rabochi.

8. Oparin, A.I., The origin of life, in The origin of life, J.D. Bernal, Editor. 1967, Weidenfeld and Nicolson: London. p. 199-234.

9. Oparin, A.I., Vozniknovenie zhizny na zemle. 1936, Moscow.: Izd. AN SSSR.

10. Oparin, A.I., The origin of life. 1938, New York: MacMillan.

11. Oparin, A.I., The origin of life. 2 ed. 1953, New York: Dover Pub. 1-269.

12. Miller, S.L., J.W. Schopf, and A. Lazcano, Oparin's ''origin of life'': Sixty years later. Journal of Molecular Evolution, 1997. 44(4): p. 351-353.

13. Miller, S.L., A Production of Amino Acids Under Possible Primitive Earth Conditions. Science, 1953. 117: p. 528-529.

14. Chyba, C.F., et al., Cometary delivery of organic molecules to the early Earth. Science, 1990. 249: p. 366-73.

15. Greenberg, J.M. and C.X. Mendoza-Gomez, The seeding of life by comets. Adv Space Res, 1992. 12(4): p. 169-80.

16. Clark, B.C., Primeval procreative comet pond. Orig Life Evol Biosph, 1988. 18(3): p. 209-38.

17. Miyakawa, S., et al., Prebiotic synthesis from CO atmospheres: implications for the origins of life. Proc Natl Acad Sci U S A, 2002. 99(23): p. 14628-31.

18. Basiuk, V.A. and R. Navarro-Gonzalez, Possible role of volcanic ash-gas clouds in the Earth's prebiotic chemistry. Orig Life Evol Biosph, 1996. 26(2): p. 173-94.

19. Keefe, A.D. and S.L. Miller, Are polyphosphates or phosphate esters prebiotic reagents? J Mol Evol, 1995. 41(6): p. 693-702.

20. Maurette, M., Carbonaceous micrometeorites and the origin of life. Orig Life Evol Biosph, 1998. 28(4-6): p. 385-412.

21. Cody, G.D., et al., Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. Science, 2000. 289(5483): p. 1337-40.

22. Wachtershauser, G., The Origin of Life and its Methodological Challenge. Journal of Theoretical Biology, 1997. 187(4): p. 483-494.

23. Leif, R.N. and B.R. Simoneit, Confined-pyrolysis as an experimental method for hydrothermal organic synthesis. Orig Life Evol Biosph, 1995. 25(5): p. 417-29.

24. Oparin, A.I., Evolution of the concepts of the origin of life, 1924-1974. Orig Life, 1976. 7(1): p. 3-8.

25. Morowitz, H.J., et al., The origin of intermediary metabolism. Proceedings of the National Academy of Sciences of the United States of America, 2000. 97(14): p. 7704-7708.

26. Morowitz, H.J., A Theory of Biochemical Organization, Metabolic Pathways, and Evolution. Complexity, 1999. 4: p. 39-53.

27. Kauffman, S.A., The origins of a connected metabolism, in The origins of order. 1993, Oxford University press: N.Y.

28. Dyson, F., Origins of Life. 1999, Cambridge: Cambridge University Press.

29. Avery, O.T., C.M. MacLeod, and M. McCarthy, Studies on the chemical nature of the substance inducing transformation of pneumococcal types. J.Exp. Med., 1944. 79: p. 137-158.

30. Lederberg, J., The transformation of genetics by DNA: an anniversary celebration of Avery, MacLeod and McCarty (1944). Genetics, 1994. 136(2): p. 423-6.

31. Segre, D., et al., Graded autocatalysis replication domain (GARD): Kinetic analysis of self-replication in mutually catalytic sets. Origins of Life and Evolution of the Biosphere, 1998. 28(4-6): p. 501-514.

32. Segre, D., Y. Pilpel, and D. Lancet, Mutual catalysis in sets of prebiotic organic molecules: evolution through computer simulated chemical kinetics. Physica A, 1998. 249(1-4): p. 558-564.

33. Segre, D. and D. Lancet, A statistical chemistry approach to the origin of life. Chemtracts - Biochemistry and Molecular Biology, 1999. 12(6): p. 382-397.

34. Segre, D. and D. Lancet, Composing life. EMBO Rep, 2000. 1(3): p. 217-22.

35. Shenhav, B., D. Segre, and D. Lancet, Mesobiotic emergence: molecular assemblies that self-replicate without biopolymers. Advances in Complex Systems, 2003. (in press).

36. Segre, D., et al., The molecular roots of compositional inheritance. Journal of Theoretical Biology, 2001. 213(3): p. 481-491.

37. Segre, D., D. Ben-Eli, and D. Lancet, Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Natl Acad Sci U S A, 2000. 97(8): p. 4112-7.

38. Shapiro, R., A Replicator Was Not Involved in the Origin of Life. Life (IUBMB), 2000(49): p. 173-176.

39. Morowitz, H.J., B. Heinz, and D.W. Deamer, The chemical logic of a minimum protocell. Origins of Life and Evolution of the Biosphere, 1988. 18: p. 281-287.

40. Segre, D., et al., The lipid world. Origins of Life and Evolution of the Biosphere, 2001. 31(1-2): p. 119-145.

41. Sargent, D.F. and R. Schwyzer, Membrane lipid phase as catalyst for peptide-receptor interactions. Proc Natl Acad Sci U S A, 1986. 83(16): p. 5774-8.