Origins: From the Protosun to the First Steps of Life
Proceedings IAU Symposium No. 345, 2018
B. G. Elmegreen, L. Viktor Tóth & M. Güdel, eds.

doi:10.1017/S1743921319002047

Early life on Earth: Tracing the chemical path from non-living to living

Addy Pross

Department of Chemistry, Ben Gurion University of the Negev, Be'er Sheva 84105, Israel
email: `pross@bgu.ac.il`

Abstract. Amongst the most tantalizing questions in science are those relating to the life issue. What is it, how did it emerge, does it exist beyond our planet? In this review some central themes that have governed this debate over past decades will be described. Through the newly proposed Persistence Principle, it is argued that material stability can be achieved through either kinetic or thermodynamic means, opening up the possibility for life to be understood as a kinetic, rather than a thermodynamic, phenomenon. That insight allows the evolutionary process from inanimate to animate to be understood as one that was initiated with the emergence of a kinetically stable prebiotic replicative chemical system. Such a chemical system, once established, was able to evolve and complexify toward increasingly stable replicative forms, toward life. With a clearer understanding of what life is, the search for life in the universe can become more clearly directed.

Keywords. Dynamic kinetic stability, origin of life, persistence principle, prebiotic chemistry.

1. Introduction

Though physicists, chemists and biologists have been addressing the origin of life problem in a scientific manner for a century now, it is remarkable how limited the progress has been. Physics and chemistry have been unable to adequately characterize the life phenomenon, with the result that biology remains a science having to address its central issues in biological terms, rather than in more fundamental physical-chemical terms. Indeed, that disconnect between physics/chemistry and biology has had a serious detrimental effect on our understanding of the origin of life (OOL) process. As a result, the question of how a prebiotic physical-chemical system could have become transformed into a biological one continues to be one of modern science's major challenges, an on-going challenge to both the physical and biological sciences.

Without a doubt the experiment that set the OOL field alight came about in 1952, when Stanley Miller, a chemistry graduate student at the University of Chicago, carried out what is arguably the most famous experiment in prebiotic chemistry (Miller 1953). Miller took a mixture of the gases thought to have made up the atmosphere of the prebiotic earth – methane, hydrogen, ammonia, and water vapor, and subjected it to an electric discharge, thereby simulating the effect of primordial lightening. The result was unexpected and dramatic; the murky mixture that resulted was found to contain a range of amino acids. The fact that amino acids, one of the key molecular building blocks of life, could be so readily formed under (presumed) prebiotic conditions seemed to signify that resolution of the OOL mystery was at hand.

Triggered by Miller's landmark experiment, several decades were spent exploring the prebiotic chemical landscape. That effort proved highly fruitful and could be thought of as part of the ahistoric approach to the problem (Fry 2000; Sutherland 2015). Plausible reactions able to produce the other main biotic building blocks - lipids, sugars, and nucleotides (from which nucleic acids are constructed), were uncovered (Sutherland 2015). Though some of these reactions, as in the case of nucleotide synthesis, required quite creative chemical thinking, it was striking that within a few decades, feasible synthetic pathways for many of life's molecular building blocks, under presumed prebiotic conditions, were uncovered. But over time it became increasingly apparent that the OOL problem is much deeper than discovering novel synthetic routes to prebiotically significant materials. In this review I will present what appears to be the essence of the OOL problem: how living things were at all able to emerge from an inanimate world, and the current status of the problem.

2. Ahistorical Nature of the OOL Riddle

Consider, you gather a group of the world's top synthetic chemists and biochemists and ask them to synthesize a simple living thing. For this synthetic exercise all required chemical materials, reagents and catalysts, whether simple or exotic, whether easy or hard to synthesize, would be supplied off the shelf, so knowledge of how those building blocks happened to appear some 3.5-4 billion years earlier would not be required. And creating the required reaction conditions? In principle, no problem there either. Any equipment able to create any desired reaction conditions would be supplied.

So how to begin? The materials chosen for the synthetic attempt would presumably be those biotic materials found within all living things - amino acids, nucleotides, sugars, lipids, and so on. But what reaction conditions should be applied? A reasonable choice would be to choose the reaction conditions that led to life's emergence some 3.5-4 billion years ago? But what were they? Over the years a variety of possible locations for life's emergence have been proposed – warm ponds, hydrothermal vents under the seas, hot springs, clay surfaces, to name key ones (Fry 2000). Worryingly, all proposals make interesting points and are thought-provoking, but none are really falsifiable, a necessary condition for a scientific hypothesis to be useful. If fact, somewhat awkwardly, we seem to have come full circle. The location debate was initiated with Darwin's 'warm little pond' (as proposed in his 1871 letter to a friend, Joseph Hooker) and has ended up in Mulkijanian's 'hot spring' (Mulkidjanian *et al.* 2012). The irony is hard to ignore. More significantly, however, none of the proposals help inform us what life is and why it emerged. The problem appears to lie elsewhere.

3. Physical Basis for Change

All physico-chemical processes are directed, and the general principle that governs that directedness is the Second Law of Thermodynamics. That's why making an omelette from eggs is easy but the doing the reverse is a lot harder. Nature's direction is from thermodynamically less stable to thermodynamically more stable. However, for the OOL process – the transition from life's building blocks to simplest life – the reaction appears to have gone the wrong way, toward the creation of a thermodynamically *unstable* system, one which depends on a continual supply of energy for that counter-thermodynamic process to be thermodynamically feasible. How could a non-equilibrium chemical system, maintained in that non-equilibrium chemical state through a continual supply of energy, have come about? Carrying out traditional chemical reactions is unlikely to lead us into this new dimension of chemical behavior. Just as exploring a 2-dimensional surface won't get you airborne and into the third dimension, undertaking traditional chemical processes is unlikely to fortuitously get you into such an energy-fueled, homeostatic non-equilibrium state. Chemists have been carrying out a wide range of chemical reactions for several hundred years now, but the kind of chemistry that leads to life-like entities does not come about. That is the reason the great physicists of the 20th century were deeply puzzled by the life phenomenon. Clearly some out-of-the-box chemical thinking would be required to get you into that very special physical-chemical state. Given these difficulties with the synthetic approach to resolving the OOL problem, let us now consider the physical/conceptual approach to the problem.

Beyond the apparent incompatibility between the Second Law of Thermodynamics and the OOL process, there is an added puzzling aspect to the OOL question. How can the inordinate complexity of all living things, even the simplest ones, be explained? Even though mathematicians and physicists are struggling to define what complexity is, as the famous quote goes, we know it when we see it. Life is incomprehensively complex (Adami 2002). Dawkins stated it succinctly in the very opening line of one of his texts: *we animals are the most complicated things in the known universe* (Dawkins 1996). Indeed, seventy years of molecular biology have brought that message home in the clearest of terms. But the reason for that complexity remains obscure. Why did simple become complex? And to muddy the waters further, not all biologists agree with the statement that complexification increased over evolutionary time. However, the realization that life's complexity emerged is itself significant, as it can open an additional means in our attempts to understanding the life process. Maybe the answer to the 'why life' question can be found by answering the 'why complexity' question. Can we somehow conceptually strip away the levels of complexity to reach some inner core, to uncover life's essence? Though life is certainly complex, could it be that the life principle is simple?

4. The Replication Reaction

Thinking along those lines, in the 1970s, Tibor Gánti, a Hungarian theoretical biologist, composed an abstract model, the Chemoton model, in which he attempted to characterize minimal life (Gánti 2003). In the model he identified living things as being composed of three sub-systems, integrated into a functional whole. His model did not mention specific molecular forms, but that is part of its appeal. It attempts to describe how any living thing would be constructed, just like an architect's drawings of a building would not need to specify the precise materials from which the building would be composed. The three components that Gánti identified, though slightly modified over the years were: (a) a self-replicating informational system, (b) a metabolic system able to supply energy and chemical building blocks, and (c) a physical compartment which encloses the cell's components. That conceptual division opened up new mechanistic horizons for experimental chemists. Reaching out for simplicity, chemists could now seek out chemical systems able to generate these capabilities separately. Indeed, through the discovery in the 1960s that certain molecules were able to self-replicate, a new direction in the origin of life studies was initiated, one whose focus was on *molecular replication.*

Replication is central to the life process, and DNA replication, as part of that general process, is key, as it is the means by which genetic information is transmitted from one generation to the next. However the process of DNA replication is a highly complex one, requiring the cellular environment for it to take place. How could such a complex multi-step process have gotten started? The first hint for how it could have taken place came in a 1967 landmark experiment, when Sol Spiegelman, a creative molecular biologist at the University of Illinois, carried out a molecular replication reaction in a test-tube (Mills *et al.* 1967). Spiegelman's experiment was conducted on an RNA molecule, a nucleic acid cousin to DNA, meaning that it is also a chain-like molecule constructed from a sequence of nucleotides. However, in contrast to DNA whose sugar component within the DNA nucleotide building blocks is deoxyribose, that in RNA nucleotides is ribose. The difference is important in that compared to DNA, which prefers to exist in a double stranded structure, the famous double helix, RNA prefers to adopt a single stranded form. That single-stranded structure means that RNA is more reactive than DNA, leading to quite different biochemical function and behavior. And that's where Spiegelman's dramatic test-tube experiment comes in.

Spiegelman discovered that mixing an RNA molecule with individual A, G, C and U nucleotides (together with an enzyme to catalyze the reaction), resulted in the RNA molecule being able to make copies of itself, much as DNA does in the cell, but in this case the process is carried out in a simple test-tube reaction. In other words RNA is a self-replicating molecule, much like DNA, but significantly, the replication reaction is able to take place outside of the highly complex environment of the living cell. Thus, to illustrate, placing an RNA molecule of a particular sequence, say AAGUCCUGAUCCUG, in a test tube with activated building blocks, A, G, U, C, and a catalyst, would result in the formation of many RNA molecules with precisely the same AAGUCCUGAUCCUG sequence.

But an even more dramatic result was to come. Spiegelman discovered that occasionally exact RNA replication did not take place. Sometimes copying errors occurred. Instead of a particular nucleotide, say G, fitting in at a certain position on the RNA template, another nucleotide, say A, would latch into place instead. As a result the RNA copy ended up being slightly different to the original. In biological parlance, a mutation had occurred. And indeed, when that replication process was carried out multiple times, by continually transferring the product of replication from one test-tube to another, an evolutionary process was revealed. Indeed, after 74 batch transfers the original RNA molecule, originally some 4000 nucleotides long, actually shortened to a fraction of its original length, just some 550 nucleotides long.

The conclusion was unambiguous: polymeric molecules with a sequence structure, as exemplified by RNA, were able, under the right conditions, to both replicate *and* evolve. That groundbreaking work was then followed up by Manfred Eigen, Peter Schuster, Leslie Orgel and others, both experimentally and theoretically (Eigen 1992). Evolution appeared to be not just a biological phenomenon, but a chemical one as well. Replicating molecules, at least certain ones, could evolve. That result appeared to resolve a long-standing dilemma in the origin of life thinking. Chemically speaking all living things are based on a dual nucleic acid-protein system. Both structurally and functionally, life's exceptional operational capabilities derive from that integrated capability. But that dual nucleic acid-protein system creates a classic 'chicken and egg' dilemma. In order for a cell to replicate, and the DNA genome within the cell to replicate as part of that general process, both nucleic acids and proteins are required. Without proteins the cell's DNA cannot be replicated, and without the coded information written into the nucleic acid sequence, the required proteins cannot be formed. Neither can come into being without the other. So how could this dual system have come into existence?

5. The RNA World

Through consideration of the Spiegelman experiment, those three grand figures, Leslie Orgel, Francis Crick and Carl Woese, offered a solution to the dilemma (Robertson & Joyce 2012). They postulated that the more complex dual DNA-protein world was preceded by a simpler world, in which RNA alone formed the basis of life at the time. RNA, being more reactive than its DNA cousin, and being present in nature as a single strand polymer, might possess some catalytic activity. Thus in that earlier world, it would have been able to carry out both replicative *and* catalytic functions, though less effectively than the dual DNA-protein system. Indeed, in the early 1980s Sidney Altman and Robert Cech confirmed experimentally that nucleic acids were able to undertake catalytic function, though, as anticipated, less effectively than proteins (Stark *et al.* 1978; Kruger *et al.* 1982). Such RNA molecules were termed *ribozymes* to express both their RNA basis and their enzymatic capability.

Reassuringly, following Max Delbruck's advice of looking carefully into cell structure to reveal the cell's history, support for the RNA-world view can be found. One key organelle, the ribosome, the exquisite molecular machine which manufactures proteins, is itself composed of several nucleic acid and protein molecules. But more careful inspection reveals an interesting feature of ribosome structure. The active core of the ribosome, where the function of generating proteins is carried out, is composed solely of RNA molecules. Those protein molecules which are present in the ribosome are only found in the ribosome periphery, not in its active core. That observation led to the conclusion that early in evolution, the primal ribosome was an RNA construction, and the incorporation of protein molecules into the structure was due to evolutionary 'fine-tuning' which came about at a later time. The existence of RNA's catalytic capabilities, together with such historic information as manifest in the ribosome structure, lent further support for the RNA-world proposal.

But that's where the good news ends. The RNA world proposal runs into a deep problem that was raised earlier. If the RNA-world thesis proposes that early life was based on RNA chemistry and was based on RNA's special chemical ability to undergo self-replication, then the study of the RNA replication reaction should throw light on the emergence of life process. At this point, however, the story becomes more problematic. The reality is that a half century of RNA study since those heady days of the 1960s have failed to reveal any indication of an evolutionary process from a simpler replicating system to a more complex and more life-like system. As discussed above, Spiegelman's RNA molecules, when subjected to continual rounds of replication, did show an evolutionary process, but toward *shorter* RNAs, i.e., away from life, not toward it. Though the extensive RNA studies of past decades have offered valuable technological benefits, no evolutionary process toward greater complexity, toward life, has been observed (Joyce 2015). Molecular evolution was directed by that ubiquitous Second Law - toward the equilibrium state, not toward life. Clearly, some crucial element was missing. In the OOL context, here was yet another dead-end.

Which leads us to the issue of metabolism, the second key component in Gánti's chemoton model (Gánti 2003). Metabolic cycles are central to the functioning of all living things. Two well-known examples are the citric acid cycle, which leads to the oxidation of acetate to carbon dioxide, and the urea cycle, which converts toxic ammonia to less harmful urea. Such cycles are not chemical cycles in the usual sense in that they are highly complex and are enzymatically controlled. The question then arises: could simple non-enzymatic cycles have appeared spontaneously on the prebiotic earth as precursors to the complex cycles of extant life? But in contrast to our discussion on replication, in which simple replication reactions are possible having been observed experimentally, the emergence of cyclic chemical organization is more problematic.

For life to have emerged from such a cycle, it would have needed to be autocatalytic. Though the term 'autocatalytic' and 'self-replicating' sound very different, they overlap to a considerable degree. The former term expresses the idea that a molecule or molecular system catalyzes its own formation, but that signifies that the system has induced a copy of itself to appear. The 'replication' term is generally used when a particular entity catalyzes its own formation directly, such as in a molecular replication reaction via a template mechanism. In contrast, the term autocatalysis is generally used for those cases where the replication comes about indirectly, through cycle formation, as is discussed below. So though the terms are quite different, in practical terms they often lead to the same result – the copying process of some entity.

In the context of the origin of life, the idea that life might have emerged from an autocatalytic cycle was first raised by Stuart Kauffman in the 1980s (Kauffman 1986). If there exists a set of molecules, or molecular aggregates, A, B, C, D, and E, and if A is able to catalyze the formation of B, B is able to catalyze the formation of C, and so on through to E, and if E is able to catalyze the formation of A, thereby closing the cycle, the cycle then becomes *autocatalytic.* Catalytic closure means the system as a whole is able to catalyze its own formation. Accordingly, if the building blocks for the formation of A, B, C, D, and E, are readily available, then the quantities of all the components within the cycle will grow exponentially. The process of autocatalysis effectively leads to a process of replication.

6. Dynamic Kinetic Stability (DKS)

As is well known, exponential growth of any kind is unsustainable. Such growth cannot be maintained as resources are quickly depleted. Thus for a replicating system to be stable over time it needs to be in a dynamic state, in a non-equilibrium steady-state, where the system's rate of formation and its rate of degradation are maintained roughly in balance. But for such a system to be maintained over time, the degraded system needs to be continually reactivated. A simple equation, which describes such a cyclic kinetic state, is shown in eq 1:

$$dX/dt = kMX - gX \tag{1}$$

where X is the replicator concentration, M is the concentration of building blocks from which X is composed (for simplicity assumed to be of one kind), and k and g are rate constants for replicator formation and decay, respectively. The kMX term represents the rate of replicator formation from its building blocks, M, while gX represents the replicator's first order rate of decay. A steady state population, a state which is effectively 'stable', can be achieved and maintained provided dX/dt remains close to zero. A simple physical metaphor to illustrate the cyclic nature of the replicative process is a water fountain. A fountain is generated when an energy source pumps water out of the fountain nozzle, after which it falls back into the reservoir, before being recycled once again (Pross 2016).

Two points should be noted. First, if the replicating system can be maintained over time, its stability is not thermodynamic, but rather kinetic – the stability kind is *dynamic kinetic stability* (DKS) (Pross 2016). The stability of the system is manifest in the stability of the population of replicators, not in the individual replicators, which are transient, as they are continually being turned over, like water in the fountain. Second, a further condition for stability is that a continual supply of energy is provided, its role being to reactivate the building blocks so that the replication reaction can be continually maintained (Pross 2016).

But in order to explain an evolutionary process for such a DKS system, an added feature is crucial. The possibility of imperfect replication – mutation – must exist. Once such mutations take place, those that are favorable in that they lead to more persistent replicators (greater DKS), will end up replacing replicators of lower stability (lower DKS). That result is implicit in the mathematics of eq 1. The direction of change – the evolutionary direction – will not be toward greater thermodynamic stability, but toward greater DKS. Of course for such a process to occur the replicating system must be structured so that it is evolvable. For nucleotide-based replicators, sequence variation indeed allows for an evolutionary process (Pross 2011).

In recent years considerable experimental support for the DKS concept has been reported. A variety of energy-fueled chemical systems, able to maintain themselves over time, have been discovered and, what's more, these are found to have very different properties to thermodynamically-controlled systems (Pascal & Pross 2015). Effectively a new dimension in chemical reactivity, a kinetic dimension, has been discovered. The discovery of that new dimension is quite dramatic in chemical terms as it offers a physico-chemical

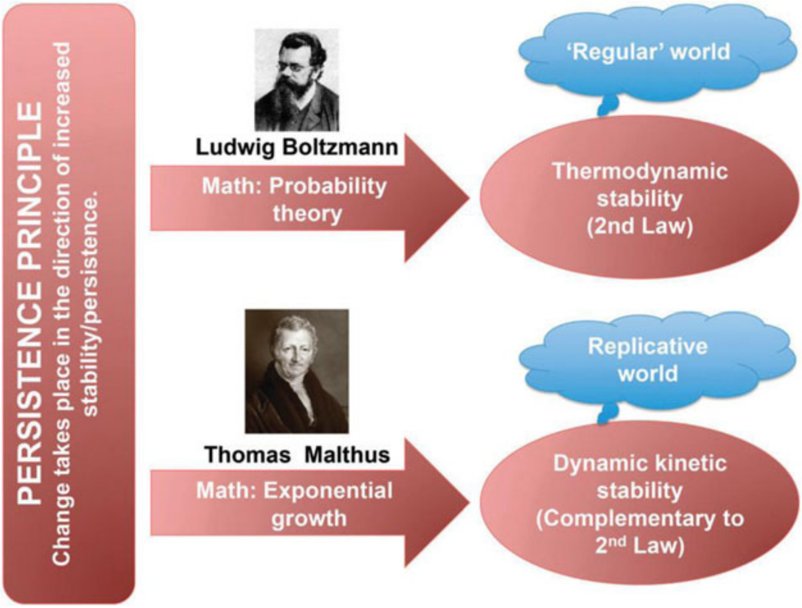

Figure 1. Schematic diagram expressing the persistence principle and the two mathematical formulations through which stability/persistence may be expressed: (a) Boltzmann's probabilistic formulation leading to thermodynamic stability, and, (b) Malthusian exponential growth leading to dynamic kinetic stability. Diagram taken from Pascal & Pross 2015.

framework for all living processes, one that has been absent till now (Pascal & Pross 2019).

7. The Persistence Principle

The realization that physico-chemical systems may become persistent for kinetic, not just thermodynamic reasons, leads to the formulation of a Persistence Principle (Pascal & Pross 2015). That principle may be stated as follows: *systems will tend from less stable/persistent to more stable/persistent forms*, or, more concisely, all matter is *driven toward more persistent forms.* The statement is logically grounded in that *less* persistent systems are, well, less persistent, meaning they will, by definition, tend to change, while *more* stable/persistent systems, again by definition, are *less* likely to undergo change. In simplest terms, changing things will change until they change into things that don't.

Of course the Second Law of Thermodynamics expresses that same directive – toward persistent forms, but through thermodynamic considerations. The Second Law seeks persistent forms through achieving the systems' equilibrium state. Significantly then, there exist two mathematical descriptions able to lead toward persistence – the Malthusian mathematics of exponential growth and Boltzmann's mathematics of statistical mechanics. Strikingly, however, each of these two mathematical kinds leads to different material consequences, with very different material expression – animate and inanimate. These two distinct paths toward greater persistence are illustrated in Fig. 1.

8. Concluding Remarks

By considering the stability concept from a kinetic perspective, rather than the more traditional thermodynamic perspective, we learn that the life phenomenon is not the inexplicable form of matter that it appears to be, one oddly incompatible with the

Second Law. A kinetic approach, based on logical/mathematical considerations, indicates that material stability can rest on an extra-thermodynamic (kinetic) base, not just on a thermodynamic one. That kinetic perspective reveals a new dimension of chemical potentiality that has only recently been discovered (Pascal & Pross 2019). That in turn leads to the new insight that life processes are but a particular expression of that kinetic chemical dimension. With a physical-chemical understanding of what life is, it is now feasible to offer a life definition: life is a self-sustained kinetic replicative network of chemical reactions whose evolutionary roots lie in an energy-fueled prebiotic replicative chemical system. The identity of that prebiotic system will likely remain unknown as the historic record of that event is long gone. But once such a replicating system emerged from the diverse environment of the early earth, the logic underpinning the persistence principle switches to one based on the mathematics of exponential growth, rather than on the equilibrium mathematics of statistical mechanics, and that switch expresses itself through the emergence of a distinctly different kind of material form - life.

As a final note, life's inordinate complexity can now be seen to derive directly from this kinetic perspective. In the functional world of replication, greater stability has been found to be enhanced by greater complexity: increasing complexity for increasing stability (Pross 2013). Accordingly, the evolutionary process leads to an increase in both stability *and* complexity. Life's mystery lies in the contingent conditions that would have enabled a primal dynamic steady-state replicative system to emerge, rather than on the nature of the evolutionary process itself. Systems chemistry laboratories are currently working on attempts to synthesize such chemical systems.

With regard the question as to whether life exists beyond our planet, the preceding discussion opens up a new approach to addressing that question. Once experiments on earth are able to outline how readily energy-fueled chemical systems able to establish dynamic replicative networks can be formed, those earth-bound experiments will have provided preliminary indications as to the likelihood of life's existence in the cosmos at large. Almost a century after Alexander Oparin, a Russian biochemist, and J.B.S. Haldane, a British evolutionary biologist, (Oparin 1952; Haldane 1929) raised the origin of life problem as one that science could potentially solve, that most imponderable of scientific questions may finally be closer to resolution.

Acknowledgements

The author is indebted to Prof. Robert Pascal for continuing discussions regarding the general role of kinetics in evolution. This publication reports the results of work undertaken within the COST actions CM1304 Emergence and Evolution of Complex Chemical Systems, and TD 1308 Origins.

References

Adami, C. 2002, *BioEssays*, 24, 1085–1094
Dawkins, R. 1996, *The blind watchmaker, Norton, NY*
Eigen, M. 1992, *Steps toward life,* Oxford UP, NY
Fry, I. 2000, *The emergence of life on earth,* Rutgers UP, NJ
Gánti, T. 2003, *The principles of life,* Oxford UP, Oxford
Haldane, J. B. S. 1929, *The Rationalist Annual* 148, 3–10
Joyce, G. F. 2015, *J. Mol. Evol.,* 81, 146–149
Kauffman, S. 1986, *J. Theor. Biol.*, 119, 1-24
Kruger, K., Grabowski, P. J., Zaug A. J., Sands, J., Gottschling D. E., & Cech, T.R. 1982, *Cell* **31,** 147-57
Miller, S. L. 1953, *Science,* 117, 528
Mills, D. R., Peterson, R. L., & Spiegelman, S. 1967, *Proc. Nat. Acad. Sci. USA* 58, 217

Mulkidjanian, A. Y., Bychkov, A. Y., Dibrova, D. V., Galperin, M. Y. & Koonin, E. V. 2012, *Proc. Nat. Acad. Sci. USA* 109, E821-E830
Oparin, A. I. 1952, *The origin of life,* Dover, NY
Pascal, R., & Pross, A. 2015, *Chem. Commun.* 5, 16160
Pascal, R., & Pross, A. 2019, *J. Syst. Chem.* 7, 1-8
Pross, A. 2011, *J. Syst. Chem.* 2, 1-14
Pross, A. 2013, *J. Mol. Evol.* 76, 185-91
Pross, A. 2016, *What is life?* 2nd Ed., Oxford UP, Oxford
Robertson, M. P. & Joyce, G. F. 2012, *The Origins of the RNA World*, in Cold Spring Harb Perspect Biol. 4, a003608
Stark B. C., Kole, R., Bowman, E. J. & Altman, S. 1978, *Proc. Nat. Acad. Sci. USA* 75, 3717-21
Sutherland, J. D. 2015, *Angew. Chem. Int. Ed.* 54, 2–20

Discussion

CUNTZ: Based on your work - what is your recommendation toward astrobiologists to optimize their efforts in the search for life in the Universe?

PROSS: To be honest - a very difficult problem. In some sense we need to progress more with what life is before we can seriously deal with seeking it elsewhere.