How life got started in our solar system

March 4, 2017

Abstract

The origin of life problem has defied explanation for more than a century. There are many different aspects to this problem, the extremely complicated biochemistry of living organism playing an important role here. I argue that the complex biochemistry is obscuring the more philosophical issues which I argue are actually rather trivial. By explaining the origin of life in terms of the relevant physical processes where the very detailed biochemistry is largely disregarded, we may get to a better broad perspective of how life came into being in our solar system.

1 Introduction

About 4 billion years ago, life emerged on Earth. While the details of how that happened are still being debated, what is clear is that living organisms are subject to the same laws of physics as all matter in the universe is. As has been pointed out by Davies and Walker, one of the fundamental problems to be solved in the origin of life problem, is to understand what life actually is [1]. They argue that it should be understood in terms of information, particularly how information is being processed at the higher level of the organism in a way that seems to defy reductionism; the naive intuition of how physics at the deeper level should explain what goes on at a higher level, seems to be violated. A living organism is a system that is going to act in a way to preserve itself, it will modify its own environment in a way that suits itself best. It will able to do that within a range of habitable environmental conditions. While we can explain how the organism works as a molecular machine, and how it gives rise to copies of itself, how you get to such systems without there being one in the first place has so far defied explanation. As pointed out in [1], the emergence of living systems from non-living systems has had to have involved a reversal of causality; in the non-living state, information at the molecular level determines what happens at higher levels, but in the living state, it's the other way around.

In this essay, I'm going to assume that there isn't anything deeply mysterious about the emergence of life anymore than the emergence of other systems such as e.g. our civilization starting from a situation where the basic building blocks of that system only existed in rudimentary forms or yet had to emerge. We may ask how our justice system came into being starting from the situation on Earth 30 million years ago. We can also ask how we are able to do our work given that none of our brain cells have the intellectual capability to understand what we are doing. Considering questions like these go a long way toward demystifying the origin of life. We don't have problems accepting that we are just complex machines and that, in general, algorithms executed by some machine do not need to exist in any of their parts but can be the result of all the parts working together. It's more challenging to explain how a machine can arise out of a system of smaller systems that start to collaborate, but even in that case we can appeal to the fact that there isn't anything deeply mysterious about our civilization gradually coming into existence starting from primitive societies. In a Chimpanzee society there exists a rudimentary form of politics [2], but unlike our society this does not form an essential backbone of society.

The real problem with explaining the origin of life is i.m.o. that the most primitive microbes we can imagine are still systems that are so enormously complex that they could not reasonably have emerged gradually from ordinary chemical processes that have no higher level of organization. This means that there must be one or or more missing link systems that are simpler than conventional living organisms but are in some sense more complex than what we would consider to be just simple chemistry. While in biology one has considered simpler systems such an an RNA world, it is also clear that such systems are still far too complex to really address the origin of life problem [3]. What we really want to know is how in a more generic sense simple molecules that interact with each other according to the laws of physics, give rise to extremely complex molecular machines that can maintain and replicate themselves. It should be clear from such an explanation that this is a natural process that will happen under the right circumstances. I'll try to get to such an explanation in the remainder of this essay.

2 Dynamic kinetic stability

Dynamic kinetic stability (DKS) is a concept introduced by Pross [4, 5, 6, 7], it describes a system of molecules that form a replicative cycle in some environment. As Pross has argued (see e.g. also the YouTube videos of part 1 and part 2 of talks given by Pross), a state where there exists a replicating cycle is fundamentally different from a state where there just exist at most catalysts that catalyzes chemical reactions but which lacks self-catalytic properties. In the latter case you'll end up at thermal equilibrium, while in the latter case the system will, after a stage of exponential growth, evolve toward becoming ever more complex due to competition from more efficient replicators. Here we note that at this stage no test-tube example of a system in the DKS state exists. All replicators that we can make are trivial, they'll just exhaust the substrate after which the system dies down into ordinary thermal equilibrium. But the theoretical properties of such systems are easy to understand, in particular that you'll end up with a chemical equivalent of Darwinian evolution.

The challenge is then to explain how one arrives at DKS starting from a situation that's best described as being close to thermal equilibrium. Here we should note that while in practice many systems we're familiar with can be described using equilibrium thermodynamics, that doesn't mean that they are close to thermal equilibrium. E.g. the chemical equilibrium state of our atmosphere would be a state where all the oxygen would have reacted with the nitrogen. But at standard conditions the reaction rate of oxygen and nitrogen is extremely small, therefore it's a good approximation to pretend that they don't react at all. So, metastability can hide the fact that the system is in fact very far from thermal equilibrium.

Metastability plays an important role in not just biology but in almost all of ordinary thermodynamics. A heat engine that exploits two heat reservoirs at different temperatures to extract work is actually exploiting the fact that the system is metastable which gives the engine the it room to bring the systems at equilibrium a bit faster. Similarly, we may consider if there are hidden metastabilties to exploit in order to get a system that is not in the DKS phase into that state. Here we need to focus on the main problem, which is that catalysts capable of catalyzing the large biomolecules that are needed to start a self-catalytic cycle are too large to arise out of chemical reactions involving only simple molecules.

3 Transitioning from ordinary chemical reactions to DKS

The problem of how large biomolecules could have arisen naturally and not just any large biomolecules but the out of the very large number of possible isomers, the few right ones, has been a major problem in biochemistry. I propose a solution that is analogous to a seemingly totally unrelated problem: How can we get to a situation where more atoms in a gas are in an excited state than in the ground state? Such a state, known as population inversion, is needed for a laser to work. The fundamental problem one faces here is that in thermal equilibrium you'll always have more atoms in the ground state than in an excited state, no matter how high the temperature is. This also means that trying to set up a non-equilibrium situation is not going to work well either, because locally the gas will still be approximately at some temperature. One can try to pump a lot of energy in a gas, but all such efforts will be in vain as long as the atoms are well described by a two level system. But suppose that there exists an excited state with an higher energy than the first excited state and we pump the atoms to such a state. If the atom in the second excited state preferentially decays to the first excited state, then we can create a population inversion if the decay time from the second to the first excited state is faster than from the first excited state to the ground state.

Similarly, the solution I propose to get to the large biomolecules needed to kick start a system in the DKS state is to consider a natural process that leads to much larger molecules than the ones we're interested in. Let's consider an environment where very cold conditions exist that can e.g. be found in proto-planets that orbit far from the Sun but which are warm in the interior. There it's possible for molecules that exists in the warm environment to diffuse into a colder environment and be deposited on some cold surface. So, layers of organic molecules may build up over time on icy surfaces. Such molecules can be stitched together under the influence of radiation. Suppose that such a proto-planet is in an elliptic orbit around the Sun, then the periodic heating and cooling will cause the layer to grow in the cold period and in the warm period chunks of ice may break off and some of the large organic molecules may end up in the warm environment. Now, the randomly stitched together molecules will not be stable in the warm environment, but large random structures will typically not break immediately into their smallest stable building blocks, as there is the possibility of a network of bonds holding a large substructure together. The fragmentation process will thus involve a hierarchical breakdown process where very large structures fragment into a few parts that then subsequently fragment into smaller and smaller parts. In the intermediary stages of this fragmentation process, molecules with catalytic properties that can impact the chemistry that goes on in the warm environment, may arise.

Note that the problem was not per se the stability of the the larger biomolecules with the right catalytic properties, rather how they would arise from smaller molecules that cannot catalyze them. So, if we get there via the fragmentation of unstable large molecules that have been cooked up in a cold region, that addresses this problem, except for the fact that the biomolecules we end up with will be some random concoction of the molecules that exist in the warm environment. So, this does not explain how the warm environment could get into a DKS phase. However, we should consider here that what matters is the long term cycle evolution of both the cold and the warm environment. As pointed out in the previous section, there may be hidden metastabilities to be exploited the presence of the two environments is such a metastability. Therefore we need to consider the evolution of both the warm and the cold environments from one cycle where molecules are deposited there, stitched together and then released into the warm environment when the proto-planet moves closer to the Sun, to the next cycle where due to the presence of the new catalysts, the chemistry in the warm environment was modified and we now end up with different molecules being deposited in the cold environment.

It is entirely plausible that the system of the warm and cold environment together could be in a DKS phase. After all, both environments are influencing each other's chemistry, albeit on the time scale it takes for the proto-planet to orbit the Sun. We are not necessarily making biomolecules that will somehow kick start the warm environment in the DKS state, it is good enough that the system as a whole is in a DKS state. Then, as is characteristic of a system in the DKS phase, one would expect that the system evolves more efficient replicative cycles. The most obvious gain is to be had by eliminating the large time scales involved in the exchange of the molecules in the different environments. Now, it has been argued by biochemists that cell membranes can easily form naturally from lipids, so one can imagine systems to arise that will function as cellular bioreactors. What is important to keep in mind here is a point made by Pross in this talk, that just getting some biomolecules or other structures essential for life to be somehow synthesized is not good enough, what matters is the entire DKS process. Obviously a collection of biomolecules and some cell walls is not alive. So, while lipids forming cell membranes that then can allow cellular life to evolve is a something that is well know, that was never the main obstacle. The main problem was always to get to the machinery of life in the first place that would then have had to move inside cells. The scenario suggested above is an plausible possibility to get to life before we get to cellular life. That this could then move inside cells so that the replicative processes then work more efficiently, can then serve as an explanation of how and why that happened.

Now, the above scenario does not provide for a clear picture of the actual biochemistry. As I already pointed out in the Introduction, while such details are important, they can cause the bigger picture to be obscured. Details such as how one gets to homochirality, i.e. that the biomolecules are either right or left handed, we don't have mixtures of both that would derail biochemistry essential for life, can be addressed by beta radiation preferentially degrading molecules of one handedness. [8]. In a DKS phase replicating cycles that rely on molecules of the wrong handedness would get eliminated due to losing the competition form its more efficient counterpart. Let's instead conclude this essay by story of how life got started on Earth, based on the above ideas.

4 Conclusion: The full story

About 4.6 billion years ago, not long after the Sun formed, dust particles started to coalesce first due to static electricity and later, larger boulders grew larger due to gravity. Many dozens of proto-planets arose in the solar system, as well as a few large gas giants. This state was unstable, the gas and dust caused friction perturbing the orbits of the gas giants that in turn would cause icy proto-planets from the outer solar system to enter the inner solar system. The plates we are familiar with today formed from a mixture of proto-planets from both the inner and outer solar system, the latter contribution is responsible for the Earth's water, the vast majority of which still resides in the Earth's interior [9]. Not all the proto-planets collided to form the planets that exist today, many of them may still exist in the Kuiper belt or beyond. Some of these entered the inner solar system in the Late Heavy Bombardment period between 4.1 billion and 3.8 billion years ago. During this period, the Lunar maria were formed due to massive impacts. One of the largest such impacts formed mare Imbrium, the impactor causing that impact had a diameter of about 250 km [10]. Impacts of such enormous size that would have hit the Earth much more frequently than the Moon, would have made the Earth a totally inhospitable place for life to get started.

But, of course, life did not start on Earth, it started on precisely on of these proto-planets that was kicked out of its orbit far away from the Sun and entered a highly elliptic orbit bringing it periodically into the inner solar system, close to the Sun. On such a proto-planet, there are warm and cold regions, large molecules can be assembled in the cold region from whatever is brewing in the warm environment. When the proto-planet comes closer to the Sun and chunks of ice melt releasing the large but very unstable molecules into the warm environment, they fragment into smaller but still large, metastable molecules that act as catalysts for reactions that would otherwise not take place. Orbit after orbit, this leads to exponential growth of certain types of molecules in both the warm and cold environments that together form an autocatalytic cycle. A growth factor of 1.05 after one orbit of, say, 6 years leads to an amplification factor of 1.55×10^{21} after just 6000 years. Due to the chemical processes lipids get formed and a slightly more efficient replicative cycle is now possible that allows for a slightly more efficient replication. But there are then just as well new possibilities that are less efficient due to the lipids. Suppose that one the new possibilities leads to a growth factor of 1.06. Then after 6000 years, this will yield a factor $13,000$ more biomolecules participating in that cycle compared to the old one. This way more efficient cycles get replaced by less efficient ones, new more efficient ways of using membranes arise, eventually giving rise to cellular lifeforms whose replicative cycle becomes divorced from the orbital cycle of the proto-planet.

When during the Late Late Heavy Bombardment period, life bearing proto-planets slammed into the Earth, Moon, and other planets, huge hunks of them broke off and veered back into space, some of these fragments would have been kilometers across [10]. Microbes are known to be able to survive inside such fragments [11].The Earth being a larger target than Venus and Mars would have been hit much more frequently than these other planets, therefore it's plausible that life arose on Mars and Venus via the transfer of such fragments before it arrived on Earth. Note that Venus is thought to have been habitable around that time [12]. Eventually, an impact hitting Mars, Venus or the Moon would cause microbe containing fragments to be scooped up by the Earth. The next proto-planet to hit Earth would wipe almost all of the life out, but note that the microbes that arrived on Earth are already quite advanced life forms that are capable of surviving tough conditions. Some them would have survived underground, shielded from the inhospitable conditions.

The explanation provided here may be criticized for not going into the biochemical details. But as I explained in the Introduction, such details would distract from the broad picture which I believe has defied the scientific community. It's similar to trying to get someone to understand how random thermal motion can be converted to macroscopic work using a heat engine. It's then far easier to explain that theoretically using thermodynamics than trying to deconstruct a real power plant in terms of all its nuts and bolts. There is then too much irrelevant information there. Now, I've also ignored most of the philosophical issues raised by Walker and Davies raised in [1] and [13]. These issues, while interesting don't have i.m.o. any bearing on actually understanding how life generically arises. It's similar to discussing how and why the universe started out in a low entropy state which ultimately does underpin the Second Law of thermodynamics. However, that issue should not distract from understanding how an heat engine works. Similarly, I believe we can get to a broad view of how life got started by separating that from these other issues.

References

[1] S.I. Walker and P.C.W. Davies, The Algorithmic Origins of Life. J. Roy. Soc. Interface **6** 20120869 .

[2] F. de Waal, Chimpanzee Politics: Power and Sex among Apes. JHU Press (2007).

[3] L. E. Orgel, Some Consequences of the RNA World Hypothesis. Origins of life and evolution of the biosphere **33**, pp 211218 (2003).

[4] A. Pross, What is life? How chemistry becomes biology? Oxford, UK: Oxford University Press (2012)

[5] A. Pross , Toward a general theory of evolution: extending Darwinian theory to inanimate matter. J. Syst. Chem. **2**, 1 (2011)

[6] A. Pross and V. Khodorkovsky, Extending the concept of kinetic stability: toward a paradigm for life. J. Phys. Org. Chem. **17**, 312316 (2004) .

[7] A. Pross, 2005 Stability in chemistry and biology: life as a kinetic state of matter. Pure Appl. Chem. **77**, 19051921 (2005).

[8] J. M. Dreiling and T. J. Gay, Chirally Sensitive Electron-Induced Molecular Breakup and the Vester-Ulbricht Hypothesis, PRL **113**, 118103 (2014).

[9] A. Morbidelli,et al. Meteoritics & Planetary Science **35**, pp. 1309-1320 (2000).

[10] P. H. Schultz and D.A. Crawford, Origin and implications of non-radial Imbrium Sculpture on the Moon, Nature **535**, 391394 (2016).

[11] N. H. Sleep and K. Zahnle, Refugia from asteroid impacts on early Mars and the early Earth, **103**, pp 2852928544 (1998).

[12] M. J. Way,et al. , Was Venus the first habitable world of our solar system?, GeophysṘes. Lett. , **43**, 83768383 (2016).

[13] S.I. Walker and P.C.W. Davies, The "Hard Problem" of Life, https://arxiv.org/abs/1606.07184.