

Hypothesis

Storage of biochemical information as the start of life: A hypothetical model for the development of the first cell

Ulrich Schreiber¹¹ University of Duisburg-Essen; ulrich.schreiber@uni-due.de

Abstract: The storage of biochemical information, which is a prerequisite for the development of the first cell, is an unsolved problem affecting all concepts of the origin of life. However, if the protected environment in the continental crust is taken into account, completely new possibilities emerge for identifying processes that may have been crucial for the formation of the first cell. Under this background, we can hypothesize that, before cellular life began, a self-sustaining cycle of molecular reaction steps with information storage in RNA existed outside of a cell. This cycle was made possible in an open system bound to gas-permeable tectonic fracture zones with a high proportion of CO₂ and N₂. The formation of peptides and vesicles in supercritical CO₂ and the chemical evolution of peptides have already been proven for the upper continental crust. Further considerations include the interactions of vesicles with catalytic peptides and the emergence of proto-tRNA. In combination with the formation of proto-tRNA synthetases, which consist of only two amino acid species and associated proto-tRNAs, the first RNA as an information storage system could have been formed with the information of proto-enzymes.

Keywords: origin of life; hydrothermal geochemistry, information storage, continental crust model, supercritical fluids, open system, binary proto-synthetase

1. Introduction

The processes underlying the origin of life remain an unsolved mystery, as the environment in which the first cell must have originated is controversial and has not been clearly identified. For this reason, it has, to date, been unfruitful to consider information storage in the first phase of life without knowing the specific starting conditions. The list of problems to be discussed in research on the origins of life is long and well known. Where did the necessary building blocks of life form in sufficient quantities and concentrate into reactive mixtures [1–5]? What were the energy sources for the relevant chemical reactions, and what processes enabled life to develop against an increase in entropy [6,7]? Due to the high thermodynamic barrier to peptide condensation in aqueous solutions, the development of the first cell in an aquatic environment is difficult to explain (paradox of water) [8–10]. If life arose on the Earth's surface, what role did solar wind and ultraviolet radiation play [11]? What effects did meteorite impacts, erosion, and sedimentation have on ponds, which must have had existed for more than millions of years, for life to develop [12]? How could vesicles develop as precursors of the cell, and how could chemical tools, consisting of a multitude of molecules with sequences that cannot be random, evolve [13,14,15]? How can the determination of the chiral molecules in terms of their present-day handedness be explained [16,17]? Last, but not least, how could information storage have started at the very beginning of life [18]? Something that is neglected in all considerations is time. One of the most difficult aspects to understand is how time has influenced the development of complex organic molecules. What is meant here by “time” is the period of hundreds of millions of years before the first cell emerged, during which, for example, entire high mountains were eroded several times down to sea level.

Until recently, there was no model for a realistic environment that could be verified in the laboratory, but such a model is a prerequisite for further discussion regarding the

evolution of the peptide/RNA system. The discussion in the literature has focused on two models in recent years: the hydrothermal model of black and white smokers (see the discussion in Russel, 2021) and warm little ponds with hydrothermal input from the crust and molecules from space [19-21]. Both models have fundamental problems. The recent white smokers were created in connection with calcareous sediments from a biogenic reef [22], a situation that cannot be transferred to the early phase of the Earth [12]. In addition, condensation reactions without supporting enzymes must have taken place in the water for polymer formation, what is considered critical [9]. The hot pond model also has several problems that call into question the origin of life in these places. In addition to exposure to ultraviolet radiation and particles from the solar wind, the first priority in relevant research is the lifespan of these systems. Shallow ponds are filled with sediment by strong weathering processes over a few thousand years. Arid regions, if they existed at all, were filled with salts. In addition, every impact of a larger meteorite, which almost exclusively occurred in the oceans, must have led to catastrophic flooding and erosion processes. Furthermore, there is no reliable estimate for the amount of organic molecules brought into the narrow catchment area by meteorites. The concentration of such molecules in meteorites is extremely low. Thus, in the short lifespans of the ponds, sufficient meteorites/micrometeorites must have struck the catchment area of the rivers to make molecules available. Meteorites with large diameters, which prevent such meteorites from burning up and the radiation in space from destroying the larger molecules within them, were determined to be capable of transporting organic molecules [23]. These molecules do not release large amounts of molecules immediately or all at once. It takes centuries to millennia for the meteorite material to weather, and the release of molecules takes place over the entire period of dissolution. Such a small amount of molecules per year cannot be considered the main resource for biochemical processes.

Within a recent discussion on the origin of the first cell in the continental crust [24-26], a realistic model was presented for the first time; this model can be tested both in the laboratory and in nature. The basis for this model is the development of the continental crust, which presumably began relatively early after the Earth had cooled down. At the point in time when the existence of the first life was postulated [27], the Earth was already more than 25% of today's mass [28-30]. Tectonic stress within the proto-crustal cores floating on a highly dynamic mantle must have led to fault zones from the beginning, which magmas and gases used as routes to the surface. Such fault zones can be active for tens to hundreds of millions of years. In addition to hot geysers in volcanic regions, there must have been numerous cooler areas in the upper continental crust due to the influx of artesian surface water, in which cold water geysers controlled by CO₂/N₂ gases were active. The conditions of a cold water geyser can be simulated in the laboratory with a high pressure cell [25,26]. Information about the typical chemistry of the fluids from the fault zones can be determined by analysing the fluid inclusions in hydrothermally formed fissure minerals. These minerals were previously found both in archaic hydrothermal quartz minerals of Western Australia and in the most recent fissure calcites in the vicinity of sub-recent volcanoes in the Eifel [30,31].

2. Environment and Resources

A multitude of different pressure and temperature conditions existed in water- and gas-filled tectonic faults in the upper continental crust [24]. Through Fischer/Tropsch-like synthesis and other processes (e.g., the Haber-Bosch process), organic molecules, which are a prerequisite for the development of life, could be formed from abiotic starting materials. Starting substances for lipids, such as aldehydes, were found in the fluid inclusions of hydrothermal archaic quartz minerals and in recent hydrothermal calcites [30,31]. Hennet et al. (1992) showed that under simulated hydrothermal conditions (150 °C, 10 Atm, pH 7) with various mineral catalysts, amino acids could be formed from the starting compounds of the fluids (formaldehyde, ammonia, cyanide) [32]. The authors analysed eight different species, of which glycine had the highest concentration. The concentrations of

other amino acids (ASP, SER, GLU, ALA, CYS, MET, and ILE) were, on average, two orders of magnitude lower. Marshall published 12 amino acid species that can be formed under hydrothermal conditions (in addition to those from Hennet et al., these species are PRO, LEU, LYS, VAL, THR, and PHE); thus, at least 14 species in the fault zones could have been available for the development of peptides [33]. However, the concentrations of the individual amino acid species must have been very different.

LaRowe and Regnier (2008) produced five common nucleobases (adenine, cytosine, guanine, thymine, and uracil) and two monosaccharides (ribose and deoxyribose) from formaldehyde and hydrocyanic acid under pressure, temperature, and bulk composition conditions that are representative of hydrothermal systems [34].

The fluids of the crustal faults have a low pH ($> 3,4$) due to high pressure when the main gas phase is CO_2 [35]. A lower pH value is achieved when sulfur compounds are present, but higher values can be achieved with a larger proportion of nitrogen in the gas composition. A short-term increase in the pH value of 0.5 is possible with a pressure drop during geyser eruption. Under high pressure and temperature and low pH conditions, marginal rocks are altered, which then, depending on the type of rock and the crustal depth, release a wide variety of metals and compounds, such as phosphate from the decomposition of apatite. Apatite occurs in high amounts in igneous rocks of the crust. The cations from the weathered minerals can act as catalysts for organic chemical reactions or can form new surfaces (sulfidic ore veins, carbonates, arsenites, and others) on the side walls. Sulphide ore contains different proportions of radioactive elements, which cause different levels of radioactive effects in chemical processes. Radioactivity can play a role in the evolution of RNA and the formation of reaction networks [36]. At the same time, clay minerals are formed and can line the rock surfaces of the fault. This fact is important when discussing the formation of RNA. Moreover, the surfaces of clay minerals can serve as templates for linking nucleotides [37,38]. Quartz is the main mineral of open fissures in tectonic faults. Fluids in the faults contain colloidal SiO_2 , which, similar to the surfaces of quartz crystals, have catalytic properties [39]. Due to the shorter distance between the moon and the Earth in the early phase of Earth's history, strong tidal forces acted on the Earth's crust [40]. While piezoelectric effects in the continental crustal rock cancel each other out during earthquakes or Earth tides due to a statistical distribution of minerals, these effects can cause a regular voltage drop and build up in the hydrothermally formed quartz in faults [41,42]. Thus, there is a cyclical weak electrical current flow that can even, as a result of earthquakes, lead to lightning-like discharges in the gas-filled cavities.

Far from influences such as solar wind, UV radiation, and erosion or flooding processes on the surface, the conditions at depth are stable for millions of years. The possible entry of molecules from Earth's surface could be facilitated by artesian water. In this way, the building blocks of life released from meteorites could reach the depths of the Earth and, for example, change the ratio of L and D amino acids [43]. The CO_2 and N_2 in mofettes and geysers are of particular importance. The ascent of CO_2 creates a current in the water column through which substances are transported. This transport ensures the replenishment of various molecules and the removal of reaction products to the surface, such that, on the one hand, no equilibrium can be established and, on the other hand, the suppression of reactions through tar formation is prevented [44]. Pure CO_2 (critical point: $30,98^\circ\text{C}$; $73,77\text{ bar}$) is already in a supercritical state (scCO_2) at a depth of approximately 1000 m (depending on the temperature and number of gas bubbles in the supernatant water column). Pure N_2 (critical point: $-146,9^\circ\text{C}$; $33,96\text{ bar}$) reaches this state at a depth of about 400 m under an open water column. Since the gases mix completely with one another, intermediate values of the critical points are formed based on the proportion of these gases. The same applies to trace gases (NH_3 , CH_4 , SO_2 , and others) if they occur in higher concentrations.

Supercritical gases have the properties of non-polar solvents, in which organic molecules can accumulate and react with one another. This interaction is favored by cavities that are filled with CO_2/N_2 at the top and water at the bottom. These cavities are autoclave-

like reaction chambers with two-phase boundaries. Here, amino acids are easily linked to peptides, with water split off. This formation is intensified by a drop in pressure, e.g., during a geyser eruption, which induces a local phase transition between scCO₂ and sub-critical gaseous CO₂ (gCO₂) at a depth of approximately 1000 m. The area affected by a phase change at this depth is approximately 100 m vertically. The resulting strong increase in entropy favours the reaction to form longer amino acid chains without catalysts [25]. It can be assumed that organic molecules that rose with scCO₂ bubbles from deeper areas of the crust accumulated here in high concentrations due to the solubility drop in gCO₂. The cyclical repetition of geyser eruptions, through which chemical evolution is significantly controlled, plays a decisive role.

3. Peptide formation

While peptide condensation is thermodynamically inhibited in an aqueous environment, the formation of peptides with up to 18 amino acid units under hydrothermal conditions was demonstrated in the laboratory in a system with excess CO₂ [25]. The conditions during the change from scCO₂ to gas favoured a condensation reaction without the need for catalysts. The main cause of this change was the enormous increase in entropy in the overall system during the transition from scCO₂ to gas.

Since it is useful to perform calculations with the smallest possible numbers of building blocks for information storage, one approach is to include only two amino acids for the initial development of functional molecules. Glycine, as shown by the experiments of Hennet et al. (1992), is the most common amino acid that can develop in a hydrothermal system (i.e., two orders of magnitude more common than ASP, SER, GLU, and ALA) [32]. The next most frequent amino acid could have been alanine or serine. While alanine hardly offers any possibilities for a more complex peptide structure, serine has a hydrophilic side chain and is, therefore, polar. In today's biology, serine is encoded by six codons, which will play a role in later considerations. If, due to the dominance of two amino acids, peptides are formed preferably from only two amino acid species, the chance for the development of a homochiral chain is relatively high. Glycine is achiral and, therefore, depends only on the second amino acid to obtain a corresponding peptide. Since the chiral amino acid is always present in equal parts for both types of handedness, a longer homochiral peptide can hardly be expected with a free combination. However, the acidic conditions of the fluids favour the formation of homochiral peptides [45]. These chains could then act as a catalyst that supports the linking of amino acids with only one type of handedness at a time. The fact that even short peptides can have catalytic functions was demonstrated by experiments carried out under different prebiotic conditions [46,47]. Notably, the catalytic activity of glycine in the peptide formation of other amino acids was also observed [48]. In this way, longer homochiral chains can be formed; these chains can be sufficient to allow the formation of secondary structures that gain stability. As noted by Danger, Plasson, and Pascal (2012), "Oligopeptides are much more stable to degradation when they adopt these specific structures, and especially if they present a β -sheet fold rather than an α -helix one" [49]. However, even short peptides have significant catalytic capabilities [50].

4. The first RNA: a precursor of tRNA?

The formation of RNA under hydrothermal conditions has not yet been proven. However, preliminary analytical results showing that phosphate and ribose or ribose and nucleobases become linked under hydrothermal conditions are very promising (oral communication by Oliver J. Schmitz, University of Duisburg-Essen). Because RNA building blocks can arise in the upper continental crust [34], the prerequisites for the formation of an RNA information storage system can be determined. This reasoning is supported by the discovery that clay mineral surfaces can act as templates for the formation of RNA [37,

38]. As a random product of chemical reactions, RNA would have no function as an information store, as RNA occurs as mRNA in later developments. The catalytic functions of ribozymes, which should have been possible at the beginning, have also been widely discussed [51-55]. The first RNA takes on a different meaning if that RNA was formed similar to a precursor to tRNA.

RNA features optimum stability under acidic conditions in the low-temperature waters of tectonic faults [56]. If longer strands are formed under such conditions (30–50 base-pairs), complementary sections accumulate and form RNA duplexes, a major problem for RNA replication. This process prevents the creation of copies through the addition of free nucleotides. Thus, the change from a single-stranded RNA template into an RNA duplex would lead to a dead-end product.

5. Thermal cycling: a simple means of separating RNA strands

Szostak (2012): “Without a simple means of separating the strands, such as thermal cycling, there is no way to continue to the next generation of replication” [57].

Faults filled with low-temperature hydrothermal water and excess CO₂/N₂ offer optimal conditions for the process of RNA copying. Temperature at a depth of less than 1000 m could have been slightly higher than 50 °C due to the ingress of surface water (artesian aquifer). The transition of scCO₂/scN₂ into the gas phase as a result of a geyser eruption causes cooling through expansion of the gas (Joule–Thompson effect). The reverse process, when the water flows back into the fault, compresses the gas, which is associated with an increase in temperature and changes to supercritical conditions. Temperatures well above 60 °C can be reached here, whereby RNA double strands become separated. With subsequent cooling until the next geyser eruption (and during the eruption), the phase of RNA replication can begin. If enough building blocks are available, strands of RNA can be copied with each eruption cycle. The prerequisite is that the low pH value is increased by a high proportion of nitrogen and decompression. These conditions are necessary to prevent protonation of the nucleotides, which occurs at low pH values. The melting temperature of DNA is significantly higher than that of RNA (higher than 80 °C), which could be one reason why DNA did not play a role in the early development of the first cells. The development of DNA as the carrier of genetic information is suspected to have occurred at a later point in time [58].

6. The formation of vesicles

The formation of vesicles and the chemical evolution of peptides in high pressure experiments were already demonstrated in previous studies [26,59], and upper crust conditions with cavities and a cold water geyser system were simulated. The focus of the experiments was the transition area from scCO₂ to gas at a depth of approximately 1000 m. The process is as follows. Along the tear-off edges, a large number of cavities are created on the fault surfaces, in which CO₂ can accumulate. These cavities provide reaction chambers with two phases—water in the lower part and scCO₂ or gCO₂ in the upper part. If scCO₂ is present, the decrease in pressure during a geyser eruption leads to a phase change in the transition region. The result is CO₂ gas in which dissolved water condenses into a mist. As experiments have shown, this process is the starting point of vesicle formation since lipids from scCO₂, which cannot remain in the gas, collect to form a primary envelope on the outer surface of the mist droplets. The sinking of the droplets to the interface between water and gas, which is also occupied by lipids, leads to a further coating, such that vesicles with a double membrane are present in the water. Additional amino acids then link to form peptides (up to 18 units) during the decrease in pressure and interact with the membranes of the vesicles formed. Mayer et al. showed that chemical evolution is possible through the cyclical repetition of pressure fluctuations, leading to the mutual stabilization of peptides and vesicles [26,59,60]. However, the large number of possible combinations in peptide formation prevents the formation of identical amino acid

chain sequences. What is missing in this process is storage of information about the sequence, as found in the RNA and DNA that formed later.

A second, less effective vesicle formation process runs parallel to mist formation in the gas. The pressure drop during a geyser eruption releases the dissolved gas in the water. Such an eruption can be so violent that, similar to opening a champagne bottle, violent turbulence and a kind of foam formation occur in the cavities. In this process, droplets that entrain lipids from the water/gas interface are catapulted into the gas space. The lipids form the first envelope around the droplets, which is then enveloped by a second envelope when the droplets fall back onto the interface. The differences between the two vesicle formation processes are obvious. While the mist droplets consist of quasi-distilled water that has forcibly absorbed organic molecules from previously supercritical gas, the "champagne droplets" are loaded with the salts of the fluid and contain organic molecules that are easily soluble in water. Another difference is the size ratio. While mist droplets are a few microns in size, the "champagne droplets" are two orders of magnitude larger.

(An estimate of the number of vesicles available for chemical evolution: Approximately 10^{11} vesicles are formed per geyser eruption. With 10,000 eruptions per year, there are 10^{15} formations, with 10^{21} over a million years. In the early phase of the Earth, the gas release was much stronger than that observed today; thus, at least 1000 cold water geysers can be expected. If the observation period is extended to 100 million years, more than 10^{26} vesicles would be available for chemical evolution.)

7. Development of life inside or outside a cell compartment?

In the discussion so far, researchers assumed that the origin of life from a certain point in time (after collecting all of the necessary building blocks) was limited to vesicles as the precursors of the cell [61]. However, all functional molecules must have developed within a narrow space considering the limited availability of the necessary building blocks. In addition to membrane proteins, which are necessary for molecule exchange and initial energy production, all other enzymes and RNA would have had to form under these restricted conditions.

Fewer explanatory problems arise if it is assumed that the laboratory of life was initially several dimensions larger. For example, the transition area from scCO_2 to gCO_2 in the fault zone can be seen as the centre of a larger unit in which reactions to form complex molecules took place. This system is supplied from below by amphiphiles, amino acids, and nucleotides formed at depth. Substances are constantly transported to the Earth's surface, some of which would block the reactions if the concentrations are too high [44]. There is unlimited potential energy from the rise of CO_2/N_2 , internal energy of the gas, thermal energy from the Earth's interior, and chemical energy from fluid-rock reactions. The accumulation of molecules from the depths leads to high concentrations in the area of the CO_2 phase change, which favours reactions to more complex units. Geyser eruptions then lead to pressure fluctuations on site and, therefore, phase transitions in which entropy increases sharply. The resulting turbulence mixes the system without the system being completely emptied. Due to gravitational separation, large molecules are preferentially retained in the vicinity of the centre. Additionally, the question of how large molecules enter the vesicles can be easily answered: The cyclical formation of vesicles with each geyser eruption is accompanied by an incorporation process of molecules in the immediate vicinity.

8. Definition of the minimum equipment needed for the start of life

To understand chemical evolution before the beginning of life, it is necessary to define the beginning of life as it could have been in its simplest form in terms of the open fault system of the crust. The start of life with a minimum number of components could be as follows:

Building blocks:

- A proto-synthetase that consists of only two amino acid species (glycine and a second amino acid (e.g., serine; valine; or an alternative amino acid, AA2)), referred to here as binary proto-synthetase.
- A second binary proto-synthetase, which also consists of the two types of amino acids (glycine and also AA2) but differs in structure from the first proto-synthetase.
- Two RNAs in which the sequences of the two binary synthetases are stored.
- Two groups of proto-tRNAs that can be specifically loaded by the synthetases.

Required environment and conditions:

- A continental tectonic fault zone with excess CO₂/N₂ (cold-water geyser system).
- Two-phase reaction chambers in the border area of the CO₂/N₂ phase change (sc/gas) at a depth of less of 1000 m with a size of <1 cm³ to > dm³.
- Unlimited supply of organic molecules (building blocks of life) from the depths.
- Ascent to the top of excess material (partially open system).

This consideration for the very first start of life assumes that the RNAs, proto-tRNAs, and synthetases were created in a previous chemical evolution process (which is explained below). The sequence of the synthetases would be stored in the RNA.

The start of life would have taken place at the moment when the cycle of the specific loading of two proto-tRNA groups by the two binary synthetases with glycine or AA2 began. The subsequent steps would represent the start of the translation process (tRNA anticodons attaching to RNAs that each store a sequence of the synthetases) and the linkage of the transported amino acid to the given binary peptide. The peptide would then assume its synthetase function by folding. For physicochemical reasons, these steps would also be possible without additional auxiliary molecules such as ribose but would be less effective and occur over a longer period of time. However, it can be speculated that the special situation in cavities with a strong increase in entropy during the CO₂ phase change would have a supporting effect. Both binary synthetases consist of the same amino acids but have different sequences. Each synthetase specifically catalyses one of its own amino acids.

The simplest form of life would therefore be the independent self-replication of different proto-tRNAs and two RNAs through a process of base pairing and the repeated copying of two specific synthetases with the help of proto-tRNAs and the information stored in the RNAs. This process would be made possible under non-equilibrium by a constant supply of components from the depths with a large excess of glycine and AA₂, as well as the special conditions in two-phase reaction chambers.

9. Did life start with a binary system?

The formation of proteins/synthetases from a minimum of two amino acids plus the development of proto-tRNAs can form the basis for a self-sustaining and multiplying molecular system, the establishment of which can be considered the start of life. Proto-synthetases are thus able to carry out the loading of proto-tRNAs. However, the exact assignment of an amino acid to a specific tRNA, as done in today's cells, was certainly not yet possible. Assuming that the four RNA bases adenine (A), cytosine (C), guanine (G), and uracil (U) were available at the beginning, 64 different combinations were accessible for the formation of anti-codons (in which there was still no definition for the stop or start codons). One promising possibility is that there were only two (or, alternatively, two groups of) proto-tRNAs out of the 64 possible proto-tRNAs to which one of the two binary proto-synthetases could carry out specific loading. In the case of two groups, the ratio of the number of proto-tRNAs between them (1:1 or otherwise) is of secondary importance. Only later, with increasing specificity, would more differentiated assignments of the proto-tRNAs to other amino acids emerge.

Ultimately, the system consists of a group of proto-tRNAs (or a single proto-tRNA) to which the amino acid glycine can be specifically loaded (0) and a second group (or a second single proto-tRNA) to which a second amino acid (AA2) can be specifically loaded (1). If serine played a role as the second amino acid, the high number of associated tRNAs still existing today may suggest an assignment to one of the two groups. The free random combination of zeros and ones (when the associated amino acids are linked) leads to any number of peptides. The anticodons can then serve as templates for complementary RNA. The simultaneous storage of the sequence via codon–anticodon interactions and the ability to read out this information again (as done later for mRNA) retain the peptides in the system for a long period of time.

This arrangement is similar to the binary system that was developed as the basis for computer technology. An 8-bit computer system defines letters or numbers with combinations of zeros and ones, which, in certain sequences, can form more complex contents (words, sentences, chapters, etc.). However, a biochemical system in the initial stage of development would not be limited to 8 bits. This system would involve rows of different lengths that allow a variety of random combinations (which can also be stored). Within this multitude of combinations, certain variations are likely to represent functional peptides, e.g., new proto-synthetases. These peptides can help increase complexity by integrating third and further amino acids with specific bonds to other proto-tRNAs (i.e., the iterative structure of an increasingly complex system). The high degree of non-specificity in this model allows a large number of tests and errors to occur without the system completely dying. An indication of the iterative structure of the chemistry of life is represented by the most frequently used amino acids of the first living organisms: glycine, alanine, valine, serine, proline, and aspartic acid [43].

10. The hypothetical model for storing information

The central part of the hypothetical model for information storage presented here is the environment of the Earth's crust at a depth of approximately 1000 m. A certain RNA type, the precursor of today's tRNA (proto-tRNA), is a prerequisite for this hypothetical model. The formation of peptides or proteins and of vesicles took place in large numbers in the cavities with each phase change from supercritical gas to subcritical gas. The mutual influence of peptides and vesicles led to chemical evolution of the peptides, which were composed of all the amino acid species available in the hydrothermal fluids. Due to physicochemical differences such as basic, acidic, polar, and non-polar, at least two groups can be postulated for this phase [see also 61]. These groups would have performed catalytic functions, each with different affinities for the amino acid species present. Moreover, some of the peptides would have had catalytic properties that led to a linkage of amino acids with proto-tRNAs. Based on the properties of the two peptide groups, at least two different groups of proto-tRNAs were loaded. It is also possible that only two different proto-tRNAs existed in this initial phase.

Due to the oversupply of glycine, this amino acid was mainly linked to one of the proto-tRNA groups. A second amino acid, which was the next most frequent or had a stronger affinity for the randomly formed peptides of the two peptide groups, was linked to the proto-tRNAs of the second group (other amino acids that were loaded onto the same proto-tRNAs appeared, but were few in number, so they only played a subordinate role). The next step was the loose coupling of loaded proto-tRNAs in a random order so that, on the one hand, the transported amino acid could be linked in a chain and, on the other hand, the anticodons could serve as templates for the formation of an RNA strand. The linkage reactions were supported by the increase in entropy during Geyser eruptions. Two amino acids comprised the majority of all formed peptides. Many of the peptides were enantiomerically pure under low pH ratios, enabling them to be folded and making them less prone to hydrolysis. At this point, the same number of peptides existed in both L and D versions.

From this repertoire (after millions or tens of millions of years) came two proteins, which, as proto-synthetases, carried out more specific loading of one of their amino acids on one of the proto-tRNA groups. The parallel formation of the associated RNA during protein production led to information in the protein blueprint being retained. Ultimately, the development of life started as described in Sections 8 and 9.

Later in the process, vesicle formation during geyser activities again became important. During each formation cycle, a large number of molecules in the vicinity of the cavities were incorporated into the vesicles. Over the course of millions of years, a sufficient number of processes took place in which proteins, tRNAs, and RNA entered the vesicles. When all necessary building blocks for the self-preservation of all molecules in a vesicle were available in large numbers (with further supply of necessary components from outside), the proto-cell could grow and divide by shearing during a turbulent geyser eruption. The still-high number of building blocks required in the two sub-cells would have enabled a new division in each case under the same conditions [12]. This process could have been the start of cellular life, which subsequently depended on the conditions of the fault zones over long periods of time. With the formation of a more complex metabolism, adaptation to new environments was possible. The conquest of the Earth's surface and the hydrothermal vents of the deep sea was only a matter of time. There are still forms of life today whose livelihoods ultimately reach back to the adaptation of prokaryotes to local energy sources [62]. Even on the continents, CO₂-bearing fault zones remain populated with archaea down to greater depths [63].

5. Conclusions

The continental crust model (CCM), which is based on gas-permeable fault zones, offers, for the first time, the possibility to carry out experiments on the formation of the first cell under realistic conditions. The physicochemical conditions of this process are known, as are the resources for the building blocks of life. Fluid inclusions from minerals formed in hydrothermal faults can be used to analyse the composition of fluid chemistry from the earliest times on earth. Moreover, the formation of vesicles in a high-pressure system of water and supercritical CO₂/N₂ is trivial. The formation of peptides occurs cyclically in connection with phase changes of CO₂ due to pressure fluctuations during geyser eruptions and Earth tides. The resulting increase in entropy enables the amino acids to be linked without the aid of catalysts. The peptides formed in this way can be divided into at least two groups. Some of these groups interact with the vesicle envelopes, leading to chemical evolution of the peptides and vesicles. The formation of proto-tRNAs can be assumed under the outlined conditions. The physical properties during pressure fluctuations lead to cyclical temperature fluctuations that allow double-stranded RNA to melt. This is the prerequisite for replication to take place. The complete range of today's tRNAs likely did not exist in the early Earth phase.

At least two different proto-tRNAs or two different proto-tRNA groups were necessary for the subsequent development of life. Since even short peptides can develop catalytic effects, albeit with low specificity, an unspecific loading of proto-tRNAs with amino acids can be postulated under the outlined conditions in an open disturbance system. The agglomeration of loaded proto-tRNAs and the simultaneous linking of amino acids with strongly increasing entropy enabled the formation of new peptides. Simultaneously using anti-codons as templates for RNA made it possible for the first time to save the sequences of the peptides. This new RNA can be seen as a precursor to mRNA (proto-mRNA). From the large number of peptides that were stored, two enzymes (binary proto-synthetases) can be postulated, both of which consisted of glycine and a second amino acid but had different catalytic properties. These enzymes were mutually able to load one of their own amino acids specifically onto one of the two groups of tRNAs. Assuming that the proto-tRNAs and the proto-mRNAs were copied in an open system and the preservation of the binary proto-synthetases was guaranteed by reading the proto-mRNA, the start of life was complete. The problem of chirality is solved by the fact that all reactions took place in

parallel with molecules featuring both types of handedness, both for the amino acids and for the RNA. As soon as the first successful life cycle with either type of handedness began, the race for handedness was decided.

Although many questions remain unanswered in this roughly sketched model, and a large number of experiments will be required to understand the individual steps in the development of life, the CCM offers a decisive advantage. The CCM is based on realistic natural environmental conditions that can be simulated in the laboratory. At the same time, there are a large number of influential factors in this environment, which, based on the corresponding variations, suggest that solutions for most problems can be found.

Among other factors, processes can be observed in this unusual environment that are difficult to implement elsewhere. For example, the question arises as to whether there can be a floatation process at suitable pH values in the fluids, leading to an accumulation of proto-tRNA on the water surface. Since tRNA has a higher density than water, ascending scCO₂ micro bubbles would have to couple with the tRNA for molecules to float to the water surface. All tRNAs have a 3' acceptor arm to which a specific amino acid is linked during loading. This arm consists of a nucleotide sequence with the bases ACC. The tRNAs with this structure are presumably the functional molecules that were extant during the earliest stages of life development. Adenine, as the outermost point of linking, is the most strongly hydrophobic base, which would enable the ACC arm to protrude into scCO₂. Loading with an amino acid during the increase in entropy as a result of the transition from a supercritical to subcritical state could take place without further catalysts. In this scenario, there would be a selection of possible non-polar, neutral amino acids for loading the proto-tRNA. These amino acids may have been leucine or valine, one of which would have to form a protein with glycine. Positioning the tRNAs in one plane directly under the water surface would make anti-codons more easily available for base pairing with the formation of RNA. In this system, due to the subsequent pressure build-up of the returning water, the temperature rises, and the tRNA melts; at the same time, the newly formed RNA leaves the anti-codons, and the amino acids are linked to form a peptide. These peptides then interact with the simultaneously formed vesicles. A chemical evolution follows, which is supported by the fact that the blueprints of the peptides are stored.

In addition, the increasingly sensitive analysis methods of trace organic chemistry offer a great opportunity to understand the developmental steps of organic molecule formation using fluid inclusions from hydrothermal quartz. Although there are no longer any quartz dikes from the oldest phase of continental crust formation, there is still a chance of finding samples from this period. The first hydrothermal quartz veins that were formed in the oldest cratons were, at some point, exposed on the surface after erosion of the overlying rock units. Weathering resulted in quartz pebbles that were transported across rivers and deposited in basins together with sand and clays. These pebbles, e.g., from the Jack Hills in Western Australia [30], are documents from the oldest continental crust that was destroyed by younger geological processes.

Funding: This research received no external funding

Acknowledgments: I am grateful to Christian Mayer, Peter Bayer, and Oliver J. Schmitz (all from the University of Duisburg-Essen) for the extensive discussions and helpful comments.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Dyson, F.J. *Origins of Life*; Cambridge University Press: Cambridge, UK, 1999.
2. Trainer, M.G. Atmospheric prebiotic chemistry and organic hazes. *Curr. Org. Chem.* **2013**, *17*, 1710–1723.
3. Airapetian, V.S.; Gloer, A.; Gronoff, G.; Hébrard, E.; Danchi, W. Prebiotic chemistry and atmospheric warming of early Earth by an active young sun. *Nat. Geosci.* **2016**, *9*, 452–455.
4. Simoneit, B.R.T. Prebiotic organic synthesis under hydrothermal conditions: An overview. *Adv. Space Res.* **2004**, *33*, 88–94.

5. Schreiber, U.; Mayer, C.; Schmitz, O.J.; Rosendahl, P.; Bronja, A.; Greule, M.; Keppler, F.; Mulder, I.; Sattler, T.; Schöler, H.F. Organic compounds in fluid inclusions of Archean quartz—Analogues of prebiotic chemistry on early Earth. *PLoS ONE* **2017**, *12*, e0177570.
6. Schrödinger, E. What is Life—the Physical Aspect of the Living Cell. *Cambridge University Press* **1944**.
7. Kleidon, A.; Lorenz, R.D. (Eds.) *Non-equilibrium thermodynamics and the production of entropy: life, earth, and beyond*. Springer Science & Business Media 2004.
8. Shapiro, Robert. Origins: A skeptic's guide to the creation of life on earth. Bantam, **1987**.
9. Marshall, M. How the first life on Earth survived its biggest threat—water. *Nature* **2020**, *588*, 210–213.
10. Russell, M.J. The “Water Problem” (sic), the Illusory Pond and Life's Submarine Emergence—A Review. *Life* **2021**, *11*, 429.
11. Cnossen, I.; Sanz-Forcada, J.; Favata, F.; Witasse, O.; Zegers, T.; Arnold, N. F. Habitat of early life: Solar X-ray and UV radiation at Earth's surface 4–3.5 billion years ago. *Journal of Geophysical Research: Planets* **2007**, *112* (E2).
12. Schreiber, U.; Mayer, C. *The First Cell - The Mystery Surrounding the Beginning of Life*. Springer 2020.
13. Deamer, D. The role of lipid membranes in life's origin. *Life* **2017**, *7*, 5.
14. Mayer, C.; Schreiber, U.; Dávila, M.J. Periodic vesicle formation in tectonic fault zones—An ideal scenario for molecular evolution. *Orig. Life Evol. Biosph.* **2015**, *45*, 139–148.
15. Mayer, C.; Schreiber, U.; Dávila, M.J. Selection of prebiotic molecules in amphiphilic environments. *Life* **2017**, *7*, 3.
16. Meierhenrich, U. *Amino acids and the asymmetry of life: caught in the act of formation*. Springer Science & Business Media, 2008.
17. Hazen, R. M.; Filley, T. R.; Goodfriend, G. A. Selective adsorption of L- and D-amino acids on calcite: Implications for biochemical homochirality. *Proceedings of the National Academy of Sciences* **2001**, *98* (10), 5487–5490.
18. Chatterjee, S.; Yadav, S. The origin of prebiotic information system in the peptide/RNA world: A simulation model of the evolution of translation and the genetic code. *Life* **2019**, *9*, 25.
19. Deamer, D.; Damer, B.; Kompanichenko, V. Hydrothermal Chemistry and the Origin of Cellular Life. *Astrobiology* **2019**, *19*, 1523–1537.
20. Damer, B.; Deamer, D. The Hot Spring Hypothesis for an Origin of Life. *Astrobiology* **2020**, *20*, 429–452.
21. Pearce, B. K.; Pudritz, R. E.; Semenov, D. A.; Henning, T. K. Origin of the RNA world: The fate of nucleobases in warm little ponds. *Proceedings of the National Academy of Sciences* **2017**, *114* (43), 11327–11332.
22. Kelley, D.S.; Früh-Green, G.L.; Karson, J.A.; Ludwig, K.A. The Lost City hydrothermal field revisited. *Oceanography* **2007**, *20* (4), 90–99.
23. Pierazzo, E.; Chyba, C.F. Amino acid survival in large cometary impacts. *Meteoritics & Planetary Science* **1999**, *34* (6), 909–918.
24. Schreiber, U.; Locker-Grütjen, O.; Mayer, C. Hypothesis: Origin of life in deep-reaching tectonic faults. *Orig. Life Evol. Biosph.* **2012**, *42*, 47–54.
25. Mayer, C.; Schreiber, U.; Dávila, M.J. Periodic vesicle formation in tectonic fault zones—An ideal scenario for molecular evolution. *Orig. Life Evol. Biosph.* **2015**, *45*, 139–148.
26. Mayer, C.; Schreiber, U.; Dávila, M.J. Selection of prebiotic molecules in amphiphilic environments. *Life* **2017**, *7*, 3.
27. Dhuime, B.; Hawkesworth, C.J.; Cawood, P.A.; Storey, C.D. A change in the geodynamics of continental growth 3 billion years ago. *Science* **2012**, *335*, 1334–1336.
28. Rozel, A.B.; Golabek, G.J.; Jain, C.; Tackley, P.J.; Gerya, T. Continental crust formation on early Earth controlled by intrusive magmatism. *Nature* **2017**, *545*, 332–333.
29. Rosas, J.C.; Korenaga, J. Rapid crustal growth and efficient crustal recycling in the early Earth: implications for Hadean and Archean geodynamics. *Earth Planet. Sci.* **2018**, *L 494*, 42–49.
30. Schreiber, U.; Mayer, C.; Schmitz, O.J.; Rosendahl, P.; Bronja, A.; Greule, M.; Keppler, F.; Mulder, I.; Sattler, T.; Schöler, H.F. Organic compounds in fluid inclusions of Archean quartz—Analogues of prebiotic chemistry on early Earth. *PLoS ONE* **2017**, *12*, e0177570.
31. Großmann, Y.; Schreiber, U.; Mayer, C.; Schmitz, O. Origin of Life: Aliphatic aldehydes in the Earth's crust—remains of prebiotic chemistry? Preprint researchsquare **2021**.
32. Hennet, R.J.C.; Holm, N.G.; Engel, M.H. Abiotic synthesis of amino acids under hydrothermal conditions and the origin of life: a perpetual phenomenon. *Naturwissenschaften* **1992**, *79*, 361–365.
33. Marshall, W.L. Hydrothermal synthesis of amino acids. *Geochim. Cosmochim. Acta* **1994**, *58*, 2099–2106.
34. LaRowe, D.E.; Regnier, P. Thermodynamic Potential for the Abiotic Synthesis of Adenine, Cytosine, Guanine, Thymine, Uracil, Ribose, and Deoxyribose in Hydrothermal Systems. *Orig. Life Evol. Biosph.* **2008**, *38*, 383.
35. Peng, C.; Crawshaw, J. P.; Maitland, G. C.; Trusler, J. M.; Vega-Maza, D. The pH of CO₂-saturated water at temperatures between 308 K and 423 K at pressures up to 15 MPa. *The Journal of Supercritical Fluids* **2013**, *82*, 129–137.
36. Yi, R.; Tran, Q. P.; Ali, S.; Yoda, I.; Adam, Z. R.; Cleaves, H.J.; Fahrenbach, A. C. A continuous reaction network that produces RNA precursors. *Proceedings of the National Academy of Sciences* **2020**, *117* (24), 13267–13274.
37. Ferris, J.P. Mineral catalysis and prebiotic synthesis: montmorillonite-catalyzed formation of RNA. *Elements* **2005**, *1*(3), 145–149.
38. Cleaves II, H.J.; Scott, A.M.; Hill, F.C.; Leszczynski, J.; Sahai, N.; Hazen, R. Mineral–organic interfacial processes: potential roles in the origins of life. *Chemical Society Reviews* **2012**, *41*(16), 5502–5525.
39. Criado-Reyes, J.; Bizzarri, B. M.; García-Ruiz, J. M.; Saladino, R.; Di Mauro, E. The role of borosilicate glass in Miller–Urey experiment. *Scientific reports* **2021**, *11* (1), 1–8.

40. Ćuk, M.; Hamilton, D.P.; Lock, S.J.; Stewart, S.T. Tidal evolution of the Moon from a high-obliquity, high-angular-momentum Earth. *Nature* **2016**, 539 (7629), 402-406.
41. Finkelstein, D.; Hill, U.S.; Powell, J.R. The piezoelectric theory of earthquake lightning. *J. Geophys. Res.* **1973**, 78, 992-993.
42. Freund, F. Charge generation and propagation in igneous rocks. *Journal of Geodynamics* **2002**, 33 (4-5), 543-570.
43. Jakschitz, T.A.; Rode, B.M. Chemical evolution from simple inorganic compounds to chiral peptides. *Chemical Society Reviews* **2012**, 41 (16), 5484-5489.
44. Benner, S.A.; Kim, H.J.; Carrigan, M.A. Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. *Accounts of chemical research* **2012**, 45(12), 2025-2034.
45. Plasson, R.; Tsuji, M.; Kamata, M.; Asakura, K. Reactivity of Alanylalanine Diastereoisomers in Neutral and Acid Aqueous Solutions: a Versatile Stereoselectivity. *Orig. Life Evol. Biosph.* **2011**, 41, 413-435.
46. Plankensteiner, K.; Righi, A.; Rode, B.M. Glycine and Diglycine as Possible Catalytic Factors in the Prebiotic Evolution of Peptides. *Orig. Life Evol. Biosph.* **2002**, 32, 225-236.
47. Sauer, F.; Haas, M.; Sydow, C.; Siegle, A.F.; Lauer, C.A.; Trapp, O. From amino acid mixtures to peptides in liquid sulphur dioxide on early Earth. *Nature Communications* **2021**, 12 (1), 1-8.
48. Jordan, I.K.; Kondrashov, F.A.; Adzhubei, I.A.; Wolf, Y.I.; Koonin, E.V.; Kondrashov, A.S.; Sunyaev, S. A universal trend of amino acid gain and loss in protein evolution. *Nature* **2005**, 433 (7026), 633-638.
49. Danger, G.; Plasson, R.; Pascal, R. Pathways for the formation and evolution of peptides in prebiotic environments. *Chemical Society Reviews* **2012**, 41 (16), 5416-5429.
50. Chatterjee, S.; Yadav, S. The origin of prebiotic information system in the peptide/RNA world: A simulation model of the evolution of translation and the genetic code. *Life* **2019**, 9, 25.
51. McGinness, K.E.; Joyce, G.F. In search of an RNA replicase ribozyme. *Chemistry & Biology* **2003**, 10 (1), 5-14.
52. Fedor, M.J.; Williamson, J.R. The catalytic diversity of RNAs. *Nature reviews Molecular cell biology* **2005**, 6 (5), 399-412.
53. Gesteland, R.F.; Cech, T.; Atkins, J.F. *The RNA World*, 3rd edn Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press 2006.
54. Chen, X.; Li, N.; Ellington, A.D. Ribozyme catalysis of metabolism in the RNA world. *Chemistry & biodiversity* **2007**, 4 (4), 633-655.
55. Lincoln, T.A.; Joyce, G.F. Self-sustained replication of an RNA enzyme. *Science* **2009**, 323 (5918), 1229-1232.
56. Järvinen, P.; Oivanen, M.; Lönnberg, H. Interconversion and phosphorester hydrolysis of 2',5'- and 3',5'-dinucleoside monophosphates: kinetics and mechanisms. *J. Org. Chem.* **1991**, 56, 5396-5401.
57. Szostak, J.W. The eightfold path to non-enzymatic RNA replication. *Journal of Systems Chemistry* **2012**, 3 (2).
58. Orgel, L.E. The origin of life on the earth. *Scientific American* **1994**, 271 (4), 76-83.
59. Mayer, C.; Schreiber, U.; Dávila, M.J.; Schmitz, O.J.; Bronja, A.; Meyer, M.; Klein, J.; Meckelmann, S.W. Molecular Evolution in 426 a Peptide-Vesicle System, *Life* **2018**, 8, 16.
60. Dávila, M.J.; Mayer, C. Membrane structure obtained in an experimental evolution. *Life* **2022**, in press
61. Carter, C.W. Coding of Class I and II aminoacyl-tRNA synthetases. *Protein Reviews* **2017**, 103-148.
62. Kelley, D.S. MANTLE. *Oceanography* **2005**, 18 (3), 32.
63. Bornemann, T.L.; Adam, P.S.; Turzynski, V.; Schreiber, U.; Figueroa-Gonzalez, P.A.; Rahlff, J.; ... Probst, A.J. Geological degassing enhances microbial metabolism in the continental subsurface. *Nature communications* **2022**, 13, 284.