Preprints.org

Hypothesis

Not peer-reviewed version

Storage of Biochemical Information as the Start of Life: A Hypothetical Model for the Development of the First Cell

Ulrich Schreiber *

Posted Date: 13 January 2023

doi: 10.20944/preprints202201.0364.v2

Keywords: origin of life; hydrothermal biochemistry; information storage; continental crust model; super-critical fluids; proto-tRNA; hydrophobicity of anti-codon

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Hypothesis

Storage of Biochemical Information as the Start of Life: A Hypothetical Model for the Development of the First Cell

Ulrich Schreiber

University of Duisburg-Essen; cogito-sum@gmx.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 cellular life began, with a self-sustaining cycle of molecular reaction steps and information storage of peptide sequences in RNA in a crustal depth of approximately 1000 m. This cycle was made possible in an open system bound to gas-permeable tectonic fracture zones with a high proportion of CO_2 and/or N_2. It can be assumed that a large number of RNA-like molecular strands formed in the crustal environment, from which a special RNA was selected by flotation processes, which formed the basis of a proto-tRNA. The formation of peptides and vesicles in supercritical CO_2 and the chemical evolution of peptides have already been proven at conditions of the upper continental crust. The behavior of individual amino acids in connection with vesicle formation deserves special interest. Hydrophobic amino acids accumulate in the vesicle membrane, with their position in the membrane depending on the degree of hydrophobicity. Selection of a proto-tRNA with an acceptor arm ending with the bases CCA resulted in a distinctive property. The position of adenine at the tip of the acceptor arm permitted it to penetrate the membrane, allowing it to be linked to an amino acid at the 3'-OH position of the terminal ribose. The penetration depth was controlled by the hydrophobicity of the opposite anti-codon, with adenine always occupying the middle position for the hydrothermally formed hydrophobic amino acids. When the vesicle membrane was closely occupied by proto-tRNAs, the anti-codons acted as templates and stored the sequence of the peptides that form. In the cases of hydrophilic anti-codons, the acceptor arm does not reach far enough into the membrane. This allowed hydrophilic amino acids to be linked to 2′-OH.

Keywords: origin of life; hydrothermal biochemistry, information storage, continental crust model, supercritical fluids, proto-tRNA, hydrophobicity of anti-codon

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

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

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 [13–15]? 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]?

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 [10]) 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 [14,15,24], 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 [25], the Earth crust was already more than 25% of today's mass [26,27]. 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 CO_2/N_2 gases were active. The conditions of a cold water geyser can be simulated in the laboratory with a high pressure cell [14,15]. 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 [5,28].

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 [5,28]. Hennet et al. (1992) showed that under simulated

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

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) [29]. The authors analysed eight different species, of which glycine had the highest concentration. The concentrations of other amino acids (ASP, SER, GLU, ALA, and ILE) were, on average, two orders of magnitude lower (CYS and MET in traces). Marshall published 12 amino acid species that can be formed under hydrothermal conditions (in addition to those from Hennet et al. [29], these species are PRO, LEU, LYS, VAL, THR, and PHE); thus, at least 12 species in the fault zones could have been available for the development of peptides [30]. 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 [31].

The fluids of the crustal faults have a low pH (> 3) due to high pressure when the main gas phase is CO_2 [32]. 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 [33]. 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 [34,35]. 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 [36]. 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 [37]. 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 [38,39]. 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 entry of molecules from Earth's surface was possible 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 [40]. The CO_2 and N_2 in mofettes and geysers are of particular importance. The ascent of gas bubbles (supercritical and subcritical) 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 [41]. Hydrophobic and amphiphilic molecules can be transported directly in the supercritical gas bubbles. If amphiphilic molecules combine with hydrophilic ones, they can be carried along with the bubbles to higher positions, like in a flotation process.

Pure CO_2 (critical point: 30,98 °C; 73,77 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 °C; 33,96 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

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 $scCO_2$ and subcritical gaseous CO_2 (gCO_2) at a depth of approximately 1000 m. The area affected by a phase change at this depth can reach more than 100 m vertically. The resulting strong increase in entropy favours the reaction to form longer amino acid chains without catalysts [14]. It can be assumed that organic molecules that rose with supercritical gas bubbles from deeper areas of the crust accumulated here in high concentrations due to the solubility drop in subcritical gas. The cyclic repetition of geyser eruptions played a decisive role in this, by which chemical evolution was significantly controlled.

3. 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 [31], 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 [34,35]. 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 [42–47]. 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 [48]. However, RNA is completely protonated at low pH values. This is not the case when scN_2 is the dominant gas. If longer strands are formed under such conditions (>20 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.

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"* [49].

Faults filled with low-temperature hydrothermal water and excess N_2 (± CO_2) 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 $scCO_2/scN_2$ 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 formation of an RNA (as a precursor to DNA) is one of the most important prerequisites for the development of an organic-chemical information system. The conditions in the fault zones with the presence of organic bases, the sugar ribose and phosphates from the dissolution of the apatites were favourable for this. However, not only the molecules that are present in today's RNA were formed in this environment. There was certainly a multitude of different sugars, which also appeared

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

in the chiral D and L versions. Alteration of the crustal rocks resulted in clay mineral lining at various sections of the fault surfaces. They probably functioned as catalysts that controlled the linking of nucleotides to form longer RNAs [34]. These RNAs could be very different due to different sugars or bases. It is therefore necessary to identify processes that, without the help of today's biochemical catalysts, led to the RNA type that had a decisive influence on the development of the first cell.

If the conditions in an environment are suitable for RNA formation, the basis for the development of an information system is not automatically given. Strands of RNA that form vary in length and are rapidly broken down into short sections by hydrolysis. If longer strands are formed, there is a high probability of segments with complementary bases in the first and last third of the RNA strand. This means that when there is a bend, the bases of both thirds are opposite each other and form a connection via base pairing. The middle third forms a loop with different diameters.

With certain strand lengths, the diameter of the loop is so small that, for reasons of space, bending is only possible if three of the bases inside the strand rotate outwards. This creates a precursor of an anti-codon to which complementary nucleotides can attach. This is almost a prototype of a tRNA (transfer RNA), which in today's cells, although much more complex, forms the mediator between a specific amino acid and an information code. For a fully formed proto-tRNA, another requirement is necessary. The end opposite the anti-codon must be single stranded. It is part of today's acceptor arm with the base sequence CCA.

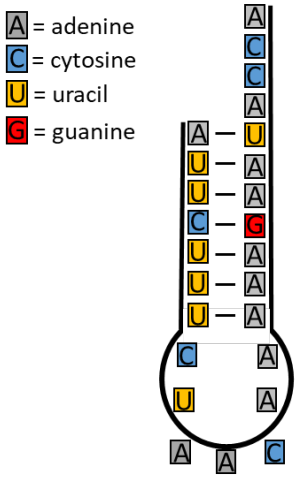

Figure 1. Example for a proto tRNA: RNA strand with formation of base pairs and an extended single strand, the last three bases of which are always CCA. Three bases are positioned in the loop (here AAC in the anti-codon; outward position requires rotation within the strand), which can be supplemented by complementary free bases. These form the codon, a triple code made up of four different RNA bases that make up RNA today.

4. The Formation of Vesicles

The formation of vesicles and the chemical evolution of peptides in high pressure experiments were already demonstrated in previous studies [15,50], and upper crust conditions with cavities and a cold water geyser system were simulated. The focus of the experiments was the transition area from $scCO_2$ 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 CO_2 can accumulate. These cavities provide reaction chambers with two phases—water in the lower part and $scCO_2$ or gCO_2 in the upper part. If $scCO_2$ is present, the decrease in pressure during a geyser eruption leads to a phase change in the transition region. The result is CO_2 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 $scCO_2$, 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 bilayer membrane are present in the water. Additional amino acids then link to form peptides during the decrease in pressure and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

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 [15,50,51]. 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.

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, intense 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 coated by a second shell 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.

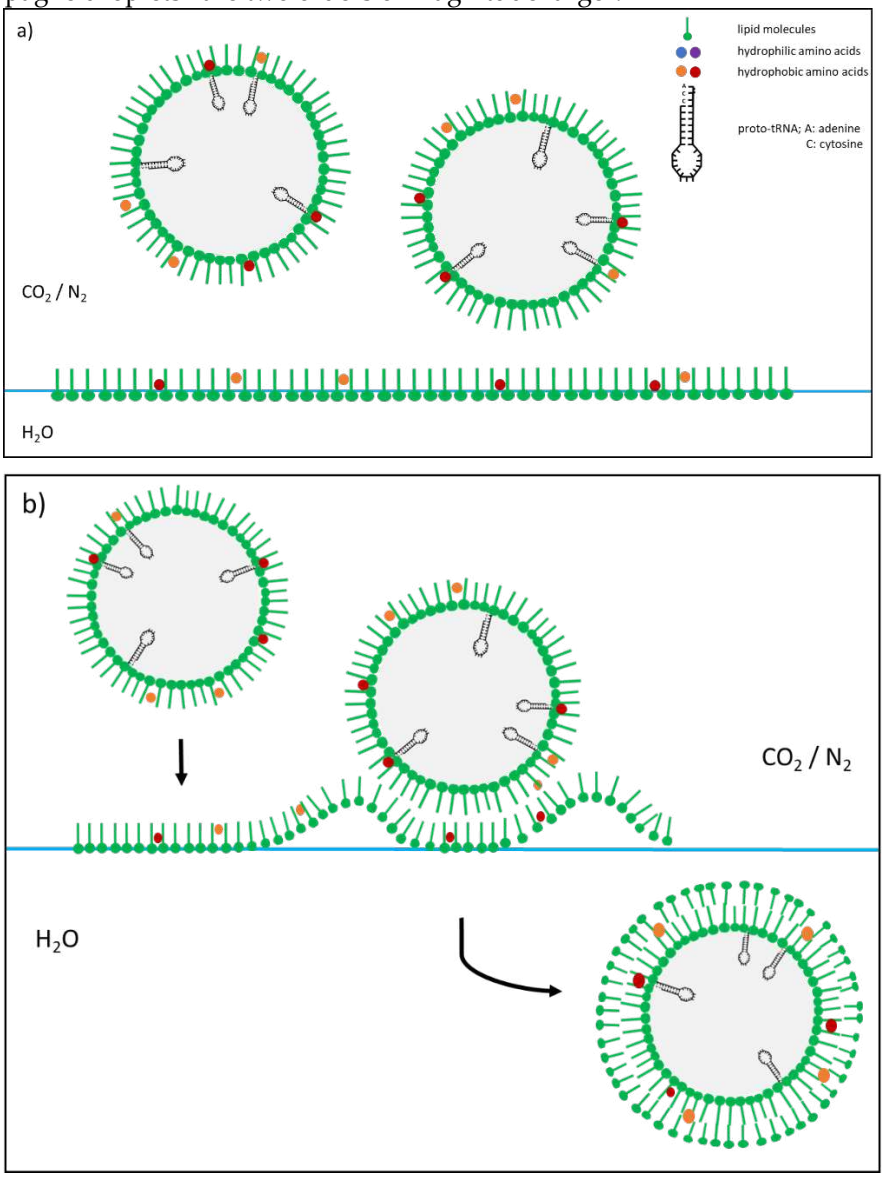

Figure 2. Vesicle formation after depressurization resulting from a cold-water geyser eruption. (**a**) Mist droplets are coated with a shell of lipids. (**b**) In contact with lipids on the water surface, a second coating is formed (based on Mayer et al. 2015 [14]).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

5. The Hypothetical Model for the Biochemical Storage of Information

A mediating molecule was required to store the information about the amino acid sequences in the peptides, which is believed to be a simple version of a tRNA. This proto-tRNA is accordingly required for the development of the first cell.

It is assumed that a large number of nucleobases, sugars and other molecules were formed in the hydrothermal environment of a fault zone and were present in different concentrations. In this case, chiral molecules were present both as L and as D enantiomers. Accordingly, there must have been a multitude of RNA strands with different lengths, conformations and molecular variations, each with sugars in L and D versions. How a certain RNA type was selected can only be recognized in retrospect after knowledge of the essential process of information storage (see below). The RNA type assumed for the hypothetical model corresponds to a precursor of today's tRNA, which combines two properties. On the one hand, an amino acid can be linked, on the other hand, the anticodon contains an information key consisting of three out of four nucleobases. For the beginning of the biochemical development, only a simply structured proto-tRNA is assumed, which had a loop with the position of the anti-codon on one side and a single strand with the sequence CCA on the other, which today corresponds to the acceptor arm. However, this alone does not result in a specific assignment of the individual amino acids to a specific anti-codon. Today, highly adapted aminoacyl-tRNA synthetases are responsible for this, which ultimately guarantee that the right amino acid is linked to exactly the right tRNA. It can only be shown that today the class I synthetase cataly zes hydrophobic amino acids at the 2′-OH end of the terminal ribose and the class II synthetase catalyzes the hydrophilic amino acids at the 3'-OH end [52,53].

Figure 3 shows how a specific RNA molecule can accumulate in a hydrothermal system. The process is comparable to flotation, in which lipids bind to organic molecules (here a proto-tRNA) (see also [54]). This creates a molecule with two different sides. The tail end of the lipid is hydrophobic, its head with the attached RNA molecule is hydrophilic. While the hydrophobic part can be trapped in supercritical gas bubbles, the hydrophilic part remains in the water. The gas bubbles rise and transport the molecules to the higher cavities and finally to the border zone, where the supercritical state changes to the subcritical one.

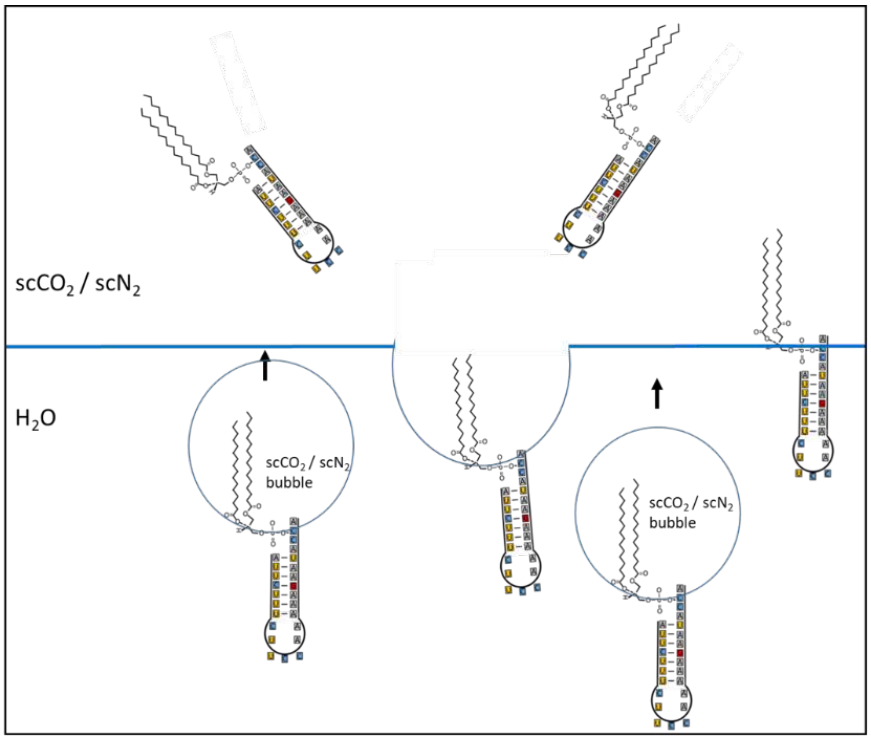

Figure 3. Principle sketch of a flotation process: Amphiphilic lipids are linked to the acceptor arm of a proto-tRNA. The hydrophobic part is integrated in bubbles of supercritical CO_2/N_2, while the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

hydrophilic head with the RNA stays in the water. The bubbles rise and collect in the roof area of a cavity. Some of the molecules get completely into the $scCO_2/N_2$ (sc = supercritical) reservoir, others position themselves exactly on the water surface. Only one RNA type is shown here. The pair formation at the tip is blocked by linking with lipid molecules (1 or more).

In this way, all RNA strands formed in the deeper crust that have a single-stranded end can be attached with one or more lipids. Transported RNAs are predominantly medium-length RNAs that have a double-stranded section and a loop. They have a random nucleotide sequence in D and L versions and the types of sugar can vary. Short strands that do not allow base pairing are rapidly hydrolyzed, longer ones are rarely formed.

In addition to many variations, there is also the combination that contains a medium length and the sugar ribose. This allows for a small loop that is spatially so tight that three bases must be rotated outwards. Other loop sizes and other sugars will not result in this configuration.

The beginning of information storage

During vesicle formation at a depth of approx. 1000 m the hydrothermally formable hydrophobic amino acids isoleucine, valine, leucine and phenylalanine can be taken up by the membrane. The hydrophilic amino acids threonine, serine, proline, glutamic acid, aspartic acid and lysine remain in the water of the droplets, which is salt-free as distilled water, but contains a large number of organic molecules. Outside the vesicles, the surrounding water of the cavities is enriched with varying concentrations of salts and dissolved inorganics. In contrast, organic molecules are less concentrated in the water outside than that of the vesicles.

On the simple proto-tRNA (Figure 1) sketched above, the base adenine is positioned with the base sequence CCA exactly at the tip of the unpaired strand (corresponds to the later acceptor arm). The acceptor arm can penetrate the membrane due to the hydrophobicity of adenine [55] and thus comes into contact with the hydrophobic amino acids (Figure 4). A linkage at the 2′-OH position of the terminal ribose in the context of the geyser cycle is now possible.

The question is how a specific connection of the corresponding amino acid in relation to the base triplet of the anti-codon could be achieved. At the base triplet, which ultimately determines the code, there are always three out of four possible bases that differ in terms of their hydrophobicity. Depending on the base and its position on the triplet, there are presumably different values for the entropic force, which in the course of the hydrophobic effect contributes to a different penetration depth of the acceptor arm into the membrane. This means that the stronger the hydrophobicity in the area of the anti-codon, the further the tip of the proto-tRNA protrudes into the membrane. There it encounters the amino acid with the highest hydrophobicity, which is most likely to be in the innermost zone. A link can be made here, whereby the amino acid binds to the 2′-OH position. Different, very finely tuned positions of the CCA arm in the membrane are achieved with appropriate combinations of bases. This allows differently hydrophobic amino acids to be linked to different positions in the membrane. If more hydrophilic bases are present at the anti-codon, CCA does not reach as far into the membrane, so that the hydrophilic amino acids are linked at the 3′-OH position.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

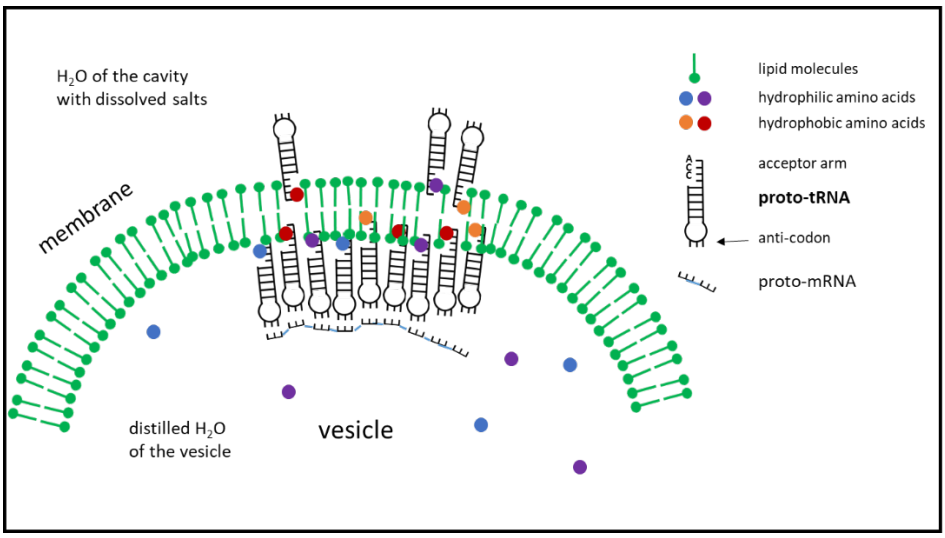

Figure 4. Penetration of the proto-tRNA into the vesicle membrane as a function of the hydrophobicity specified by the anti-codon. After the amino acids have been linked to the proto-tRNA, a peptide is formed by linking the amino acids to one another while at the same time separating them from the proto-tRNAs. The tightly packed anti-codons form the information store for the sequence of the peptide and serve as a template for a proto-mRNA, which is created by attaching complementary nucleotides and is then separated (only molecules relevant to the process are shown here).

If the terminal ribose is at the level of the lipid heads on the inner edge of the membrane, there is no loading with an amino acid. The positions here correspond to the stop codons. In Table 1, when comparing the anti-codon assignments with the associated amino acids, it becomes clear that all hydrothermally formable hydrophobic amino acids have an adenine in the middle position and some additionally in the edge positions. The hydrophilic amino acids have a correspondingly opposite occupancy. Weber and Lacey [56] already discovered this striking correlation without being able to offer an explanation for it.

Table 1. The canonical amino acids with decreasing hydrophobicity (Kyte and Doolittle scale [57]). Those that can be formed hydrothermally are highlighted in colour. The anti-codons of the associated tRNA from the hydrophobic (hydrothermal) amino acids have the hydrophobic adenine (in red) in the middle position (each second letter).

Kyte and Doolittle scale			Variations	Codon	Anti-Codon
				1-2-3 (position)	1-2-3 (position)
I	Ile	Isoleucine	3	AU ACU	UA UGA
V	Val	Valine	4	GU GACU	CA CUGA
L	Leu	Leucine	6	UU AG - CU UCAG	AA UC - GA AGUC
F	Phe	Phenylalanine	2	UU UC	AA AC
C	Cys	Cysteine	2	UG UC	AC AG
M	Met	Methionine	1	AU G	UA C
A	Ala	Alanine	4	GC GACU	CG CUGA
G	Gly	Glycine	4	GG UCAG	CC AGUC
T	Thr	Threonine	4	AC GACU	UG CUGA
S	Ser	Serine	6	AG CU - UC UCAG	UC GA - AG AGUC
W	Trp	Tryptophan	1	UG G	AC C
Y	Tyr	Tyrosine	2	UA UC	AU AC
P	Pro	Proline	4	CC GACU	GG CUGA
H	His	Histidine	2	CA UC	GU AG
E	Glu	Glutamic acid	2	GA GA	CU CU
Q	Gln	Glutamine	2	CA AG	GU UC
D	Asp	Aspartic acid	2	GA CU	CU GA
N	Asn	Asparagine	2	AA CU	UU GA
K	Lys	Lysine	2	AA GA	UU CU
R	Arg	Arginine	2	AG GA	UC CU
Stop	CCA position in lipid head region?		2	UA AG	AU UC
Stop	CCA position in lipid head region?		1	UG A	AC U

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

If the interior of the vesicle membrane is densely occupied by the proto-tRNAs, the loops with the anti-codons are close together. This stabilizes the molecular grouping as a whole. This offers the possibility of forming templates to which free, complementary nucleotides or already shorter RNA sections can attach. In the event of further pressure fluctuations, the amino acids can be linked to form peptides, while the RNA formation (proto-mRNA) ensures the storage of the sequence via the templates.

The conditions on the outer shell do not offer this possibility due to the convex curvature. Penetration into the membrane and the linking of the proto-tRNAs with the amino acids can take place in a similar way. But the loops with the anti-codons are at a greater distance from one another, even if they are close together, so that no template can form. If these individual proto-tRNAs loaded with amino acids are included in the vesicles during the next geyser eruptions, they are available for loading the proto-mRNA inside the vesicles and thus for the copying process. The same applies to the area outside of the cells in the water of the cavities into which the proto-mRNAs end up after the cells have been destroyed.

The moment when the first enantiomerically pure peptide with a catalytic function appeared, the sequence of which was stored in an RNA and preserved by copying, this can be regarded as the beginning of life. The handedness of this starting peptide determined the orientation obtained so far.

The start was complete with the storage of the sequence of a catalytically active peptide and its propagation. The further steps consisted in the continuous development of new peptides (while at the same time storing the sequences), from which millions of years arose again those that had catalytic functions and helped to support reactions that maintained the system "vesicle formation". From a certain point in time, biochemical selection resulted in the cells being equipped in such a way that, with the supply of external components, all the necessary building blocks could be multiplied. After the single proto-cells had grown sufficiently, physical forces (shearing during flow turbulence) were sufficient to cause the cells to divide. The daughter cells could continue to develop if there was a surplus of building blocks in the mother cell beforehand, which were still present in sufficient numbers in both daughter cells after the division. Already here there was a large number of starter cells, all of which had different (quantitative) equipment, but possessed the same "language and grammar".

The hypothetical development path to the first cell is shown in a very simplified manner. In all development steps, it must be taken into account that a large number of molecules existed within a group. Accordingly, there was a diverse range of competitors, for example RNAs within the vesicles.

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.; Glocer, 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**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

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. 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.
26. 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.
27. 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.
28. Großmann, Y.; Schreiber, U.; Mayer, C.: Schmitz, O. J. Aliphatic Aldehydes in the Earth's Crust—Remains of Prebiotic Chemistry? *Life* **2022**, 12(7), 925.
29. 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
30. Marshall, W.L. Hydrothermal synthesis of amino acids. *Geochim. Cosmochim. Acta* **1994**, 58, 2099–2106.
31. 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.
32. 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.
33. 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.
34. Ferris, J.P. Mineral catalysis and prebiotic synthesis: montmorillonite-catalyzed formation of RNA. Elements **2005**, 1(3), 145-149.
35. 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.
36. 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.
37. 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.
38. Finkelstein, D.; Hill, U.S.; Powell, J.R. The piezoelectric theory of earthquake lightning. *J. Geophys. Res.* **1973**, 78, 992–993.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 January 2023 doi:10.20944/preprints202201.0364.v2

39. Freund, F. Charge generation and propagation in igneous rocks. *Journal of Geodynamics* **2002**, 33 (4-5), 543-570.
40. Jakschitz, T.A.; Rode, B.M. Chemical evolution from simple inorganic compounds to chiral peptides. Chemical Society Reviews **2012**, 41 (16), 5484-5489.
41. 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.
42. 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.
43. McGinness, K.E.; Joyce, G.F. In search of an RNA replicase ribozyme. *Chemistry & Biology* **2003**, 10 (1), 5-14.
44. Fedor, M.J.; Williamson, J.R. The catalytic diversity of RNAs. *Nature reviews Molecular cell biology* **2005**, 6 (5), 399-412.
45. Gesteland, R.F.; Cech, T.; Atkins, J.F. *The RNA World*, 3rd edn Cold Spring Harbor. NY: Cold Spring Harbor Laboratory Press 2006.
46. Chen, X.; Li, N.; Ellington, A.D. Ribozyme catalysis of metabolism in the RNA world. *Chemistry & biodiversity* **2007**, *4* (4), 633-655.
47. Lincoln, T.A.;Joyce, G.F. Self-sustained replication of an RNA enzyme. *Science* **2009**, 323 (5918), 1229-1232.
48. 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.
49. Szostak, J.W. The eightfold path to non-enzymatic RNA replication. *Journal of Systems Chemistry* **2012**, 3 (2).
50. 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.
51. Dávila, M. J.; Mayer, C. Membrane structure obtained in an experimental evolution process. *Life* **2022**, 12(2), 145.
52. Carter, C.W. Coding of Class I and II aminoacyl-tRNA synthetases. *Protein Reviews* **2017**, 103-148.
53. Eriani, G.; Delarue, M.; Poch, O.; Gangloff, J.; Moras, D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **1990**, 347 (6289), 203-206.
54. Baillet, J.; Desvergnes, V.; Hamoud, A.; Latxague, L.; Barthélémy, P. Lipid and nucleic acid chemistries: combining the best of both worlds to construct advanced biomaterials. Advanced Materials 2018, 30 (11), 1705078.
55. Shih, P.; Pedersen, L. G.; Gibbs, P. R.; Wolfenden, R. Hydrophobicities of the nucleic acid bases: distribution coefficients from water to cyclohexane. *Journal of molecular biology* **1998**, 280 (3), 421-430.
56. Weber, A. L.; Lacey, J. C. Genetic code correlations: amino acids and their anticodon nucleotides. *Journal of Molecular Evolution* **1978**, 11 (3), 199-210.
57. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **1982**, 157 (1), 105-132.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.