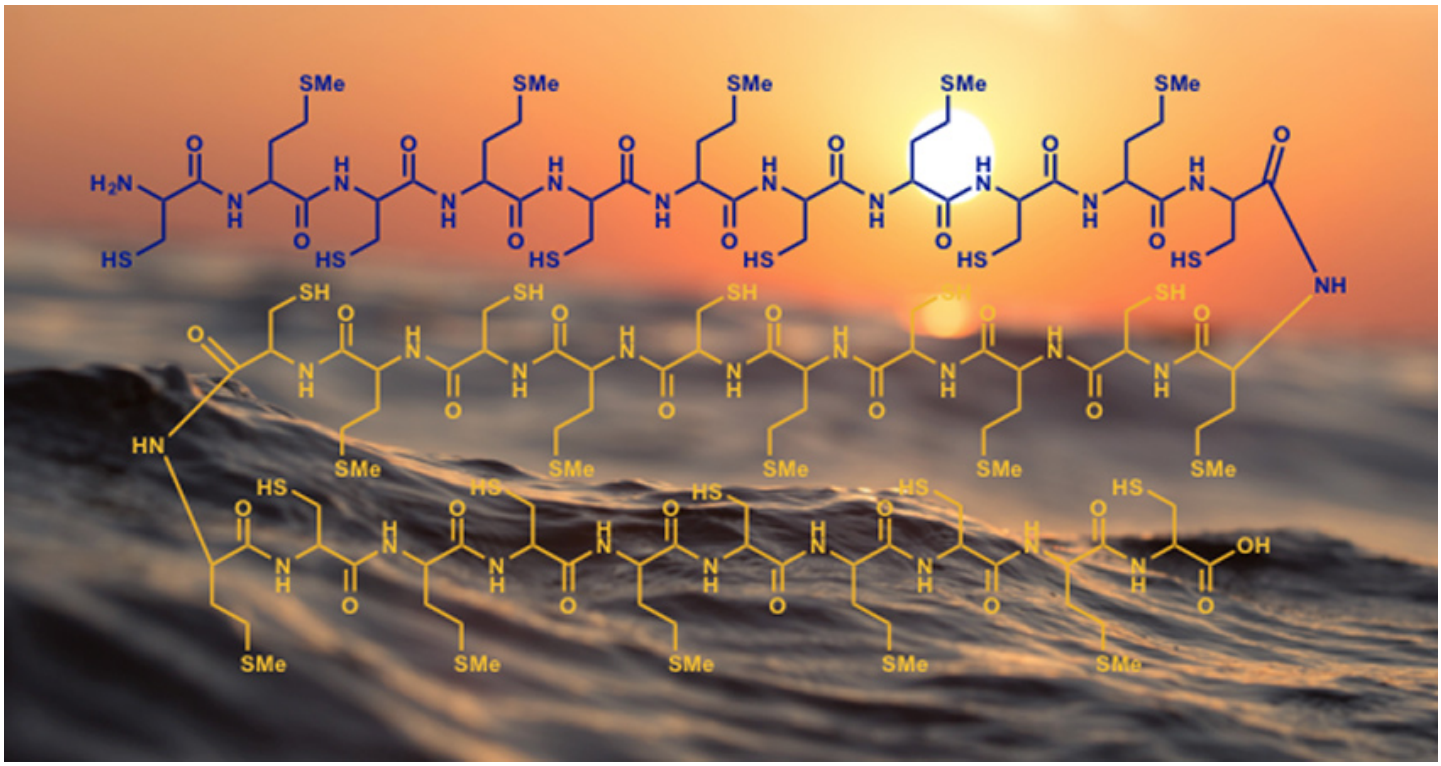
Once upon a time when life appeared: chemistry in the earth's ocean 4 billion years ago

Auteur :
VALLÉE Yannick, Professeur de Chimie, Département de Chimie Moléculaire, Université Grenoble Alpes

02-04-2019



Some 4 billion years ago, the Earth was largely covered by a huge ocean. This ocean contained a large number of small organic molecules, which are called "prebiotics" because they were there before life appeared. What were they? Were they synthesized on site or did they come from space? How did they bind to form long polymers, some of them carrying genetic information, others working to reproduce all the basic molecules and then aggregate them into polymers? Nothing was certain in advance! The polymers were not stable, the bonds were difficult to create. Yet here we are. It is therefore good that a chemistry of extreme subtlety found the energy and time necessary to set itself up. This article gives some hints to try to understand what may have happened.

1. A long time ago, in a galaxy not so far away

About 4.6 billion years ago, in an off-centre location in the spiral galaxy we call the Milky Way, a vast disc of matter was formed. Most of the gases, grains, blocks that made up this disc have concentrated and fused, to form a star, the Sun. The little amount of residual matter that remained formed planets and smaller objects, dwarf planets and asteroids. Our Earth, in fact a proto-Earth, was formed at that time. Fifty million years later (not much compared to the scale of astronomical time) this proto-Earth was struck by a very massive object, Theia, a planetoid{ind-text}Small celestial body with some characteristics of a planet. The term refers to structures as varied as asteroids, dwarf planets, protoplanets, etc. {end-tooltip} the size of Mars. From this gigantic shock came the Moon and our present Earth [1].

Before this shock, it is likely that the Earth's atmosphere contained a lot of hydrogen, the major component of the proto-solar disk. But the shock was huge. The light elements were expelled. A new atmosphere resulted, rich in carbon dioxide (CO_2), nitrogen (N_2) and water vapour (H_2O). The Earth was still very hot. However, it cools down quite quickly. The water vapour condensed and fell in an incessant torrential rain to form **a first, unique and immense ocean**.

Under this ocean, the upper part of the Earth's mantle solidified, forming a first crust. The first tectonic plates gradually took place. Perhaps emerging from the ocean, we could already see some proto-continents, scattered islands, probably **volcanoes much more active than our current volcanoes**. Our planet was still full of energy! This compensated the weakness of the young Sun, which was smaller and less powerful than today. Without the energy released by the planet, without the significant greenhouse effect resulting from the high proportion of CO_2 in the atmosphere, it is quite possible that all the water has become ice. What kind of life could have been born in this ice? No doubt, none...

Just as the proto-Earth crossed Theia's path, after the birth of the Moon, the Earth was struck by many asteroids that probably brought it a good amount of extra water, perhaps also organic molecules. After an intense rebound from these cataclysmic bombardments, they ended **3.8 billion years** ago (well, almost ended: we are still not immune to a catastrophic shock. Dinosaurs would not say the opposite!).

It is not impossible that life appeared before the end of this "great late bombardment{ind-text}Period in the history of the solar system extending approximately 4.1 to 3.9 billion years ago, during which there was a significant increase in meteoric or cometary impacts on the telluric planets.{end-tooltip}", but the evidence in this regard remains tenuous. And even if it had started, would it have survived these repeated disasters? (see The origin of life as seen by a geologist who loves astronomy).

2. So much water! So much water!

So let us place ourselves, a little less than 4 billion years ago, at the end of a geological era called the Hadean. At that time, the Earth had a gigantic ocean, hyperactive volcanoes, embryos of continents. The Moon was moving away from it, but it was far from reaching its current orbit: it was still three times closer. As a result, the force of the tides was gigantic, more than twenty times greater than it is today. The winds were impressive. Even if it was cooling down, the **ocean temperature was probably warmer than it is today**.

It is difficult to know its pH{ind-text}Abreviation for Hydrogen Potential, a measure of the activity of the hydrogen ion (or proton) in a solution. The pH is an indicator of the acidity (pH below 7) or alkalinity (pH above 7) of a solution. A solution of pH 7 is called neutral{end-tooltip}: it is currently slightly basic (around 8.1) but is progressively acidified because of human CO_2 emissions. In water, CO_2 forms an acid: carbonic acid. There is much less CO_2 in the atmosphere today than when life was born. Perhaps the ocean was therefore rather acidic at the time, which had an impact on the chemistry that could occur there. Of course, it contained ions. As today sodium (Na^+) and chlorides (Cl^-) dominated. **The ocean was already salty!** There was also calcium, magnesium, bromides, and even much more iodides than today.

At first it was thought that the primitive atmosphere was very reductive, that it contained a lot of hydrogen, methane, ammonia. But we saw it, if there had been hydrogen, the shock with Theia caused the expulsion of this light gas into space. However, the atmosphere was not oxidizing. It contained very little or no molecular oxygen (O_2). We can date the appearance of a significant amount of O_2 on Earth by determining the age of the oldest deposits containing ferric iron. Indeed, when iron is exposed to water containing oxygen, it rusts. That is to say, it is oxidized to ferric iron (Fe^{3+}). Ferric iron is not soluble in water. Without oxygen, however, iron forms ferrous ions (Fe^{2+}) which are soluble.

There is therefore a major difference between our present ocean and **the primitive ocean: the latter contained dissolved ferrous iron**, not ours.

3. A fullness of small molecules

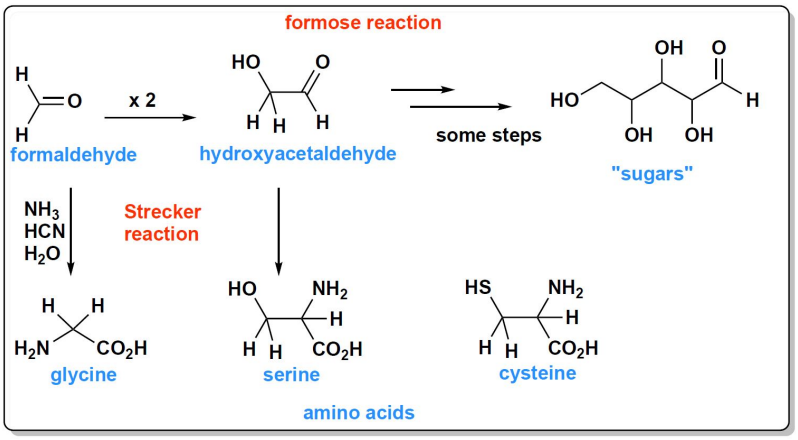


Figure 1. Examples of prebiotic syntheses. The "sugar" represented, a pentose, with 5 carbon atoms, is only one example among many. Similarly, it is possible to consider the synthesis of other amino acids than glycine, serine and cysteine.

This ocean also contained organic molecules. From CO_2 or methane (CH_4), molecules with two carbon atoms are easily formed. CO_2 can be reduced to **formaldehyde** (H_2CO, Figure 1) which by a reaction called "formose reaction{ind-text}Word formed by the contraction of the terms formaldehyde and aldose. This reaction, discovered by the Russian chemist Alexander Boutlerov in 1861, consists in polymerizing formaldehyde to form sugars including pentoses (sugars with five carbon atoms). This reaction is important in the abiotic formation processes of living molecules. {end-tooltip}" first gives hydroxyacetaldehyde{ind-text}Molecule of chemical structure $C_2H_4O_2$. It is the simplest molecule that has both a hydroxyl group (OH) and an aldehyde group (CHO).{end-tooltip} (a molecule with two carbons) and then longer molecules that are sugars. A cook would say that with these sugars in addition to sodium chloride, the Hadaean ocean was sweet and sour!

In addition to sugars, at least two types of molecules are needed to build a living cell: proteins and nucleic acids. These molecules all contain nitrogen. What could be the sources of this nitrogen in the prebiotic ocean? Probably ammonia (NH_3), and hydrocyanic acid (HCN). When these two compounds react with formaldehyde they give the simplest amino acid, **glycine**. This molecule is synthesized through an essential reaction called *Strecker* synthesis, named after Adolph Strecker, a Germanic chemist, who discovered it in the mid-19th century. This synthesis, from other aldehydes, can give various amino acids, for example **serine** from hydroxyacetaldehyde (Figure 1).

These amino acids are the basic components of the polymers that are proteins. As we have just seen, they could be synthesized quite easily: they were therefore most likely present in the early ocean.

What about nucleic acid precursors, DNA{ind-text}A abbreviation for deoxyribonucleic acid. DNA is a macromolecule composed of nucleotide monomers formed of a nitrogenous base (adenine, cytosine, guanine or thymine) linked to deoxyribose, itself linked to a phosphate group. It is a nucleic acid, like ribonucleic acid (RNA). Present in all cells and in many viruses, DNA contains genetic information, called genome, that enables the development, functioning and reproduction of living beings. The DNA molecules of living cells are formed by two antiparallel strands wrapped around each other to form a double helix.RNA is a macromolecule consisting of a sequence of ribonucleotides (adenine, cytosine, guanine, guanine, uracil) linked together by nucleotide bonds and performing many functions within the cell. It is a nucleic acid, like DNA. {end-tooltip}, which carries genetic information? Their synthesis is a little more complex. It is less obvious that all of them were present. But it is possible to write prebiotic syntheses for each of them. The ribose{ind-text}Ribose is a ose (sugar) made up of a chain of five carbon elements and an aldehyde function. It is a component of RNA used in genetic transcription. It is related to deoxyribose, which is a component of DNA. It is also present in many molecules important in metabolic processes (in particular ATP or adenosine triphosphate). {end-tooltip} can thus be obtained by the formose reaction already mentioned, nucleic bases from hydrogen cyanide, and direct access paths to ribose-base complexes have recently been published by researchers.

The synthesis of DNA and RNA chains nevertheless raises the question of the **source of phosphorus**. It is indeed abundantly present in these polymers that carry genetic information. In our world oxidant{ind-text}In chemistry, a chemical element is oxidizing when it gives one or more electrons during an oxidation-reduction reaction (see also oxidation-reduction and reductant in the glossary).{end-tooltip} world, this element is generally found in the form of phosphate, especially calcium phosphate, insoluble in water. Were there phosphates (soluble) in the primitive non-oxidizing ocean? If not, what was the source of soluble phosphorus? This is an open question [2].

Another important element is the **sulphur** present today in two life-sustaining amino acids, methionine and cysteine. It is released in significant amounts from active volcanoes, fumaroles, many hydrothermal springs, often in the form of hydrogen sulfide (H_2S). It is therefore reasonable to assume that the primitive ocean contained hydrogen sulphide, and consequently small sulphur molecules, such as the amino acid cysteine.

If we are certain (some would say: almost certain) that no little green man from intergalactic space has ever set foot on our planet, it is not impossible that some of the molecules mentioned here have landed on Earth, brought by the millions of asteroids that struck it, especially during the Great Late Bombardment. Thus, the magnificent visit of the Rosetta probe to the asteroid 67P/Chourioumov-Guérassimenko, known as "Chouri", showed that it contained water, ammonia, formaldehyde, hydrocyanic acid, hydrogen sulphide... but also more complex organic molecules including glycine, this small amino acid of which we have described above a possible synthesis on Earth (Read How to study the organic molecules of comets?).

So the first glycine: "terrestrial" or "extraterrestrial"? What about the other amino acids? What about the bases of DNA? No one knows it. But what is certain is that when the massive bombardments stopped, when this possible alien source dried up, when all these alien molecules were used, terrestrial syntheses had to take over. As we have seen, they are quite possible. The extraterrestrial hypothesis, if it cannot be refuted, is not essential to describe the birth of life on Earth.

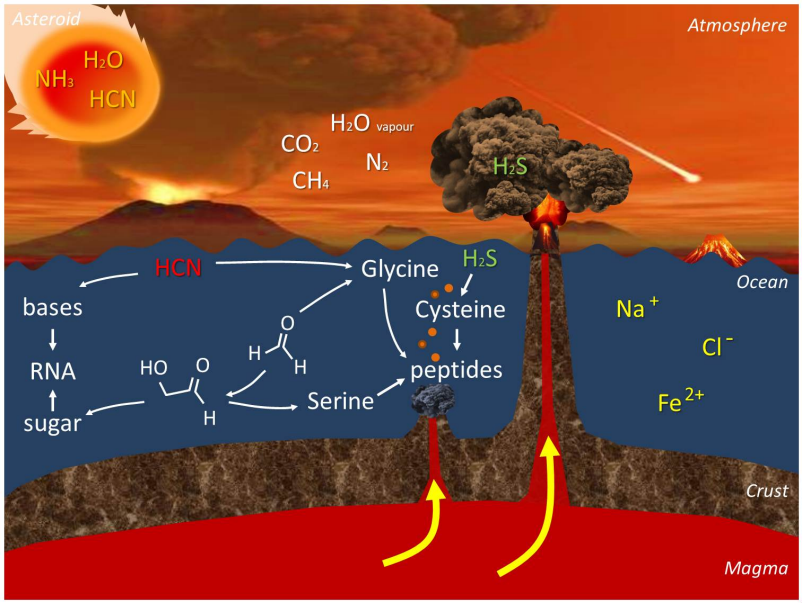


Figure 2. 4 billion years ago, a first Earth crust was formed, largely covered by a vast salty ocean containing soluble ferrous iron. Asteroids brought water and small organic molecules. Other molecules were formed in the ocean. The presence of hydrocyanic acid HCN then allowed on the one hand the synthesis of RNA bases, on the other hand that of amino acids which, upon polymeriziation, gave first peptides.

Figure 2 summarizes all this chemistry.

There remains the problem of the concentration of these molecules. This is a very important question: the more diluted the starting compounds of a given reaction are, the slower the reaction is. Certainly life had time ahead of it. However many products resulting from the condensation of small molecules are not very stable in water. We must be able to do enough, fast enough, so that they continue to grow, to form longer and longer molecules, more and more complex, before separating again into small starting components. This raises two questions: How much water was there on Earth? And this water, what mass of organic molecules did it contain?

How much water? The most realistic hypothesis is to consider that there were, roughly speaking, no more or less than today, *i.e.* about **1.36 billion km^3**, *i.e.* by rounding off a thousand billion billion litres, which is not insignificant!

Knowing how many organic molecules there were on the way to life is much more complicated. The current terrestrial biosphere contains 2,000 gigatonnes (2,000 billion tonnes = 2 billion billion grams) of organic carbon. It is very unlikely that there were more, in the form of small molecules, at a time when precisely "organic" life had not yet appeared.

Let's do the math: 2 billion billion billion grams divided by 1,000 billion billion litres, that's 2 milligrams of organic carbon per litre of water. It is a low concentration, but it is not totally ridiculous. It is probable that this leads to an overestimated idea of the concentration of organic molecules in the Hadaean ocean. The actual concentration was probably even lower. So? How can we envisage fairly rapid reactions in this ocean, continuously stirred by gigantic tides, and therefore roughly homogeneous?

Darwin himself had already expressed the problem when he wrote to his friend Joseph Hooker in 1871:" *"But if (and oh what a big if) we could conceive in some warm little pond with all sorts of ammonia and phosphoric salts, light, heat, electricity etcetera present, that a protein compound was chemically formed, ready to undergo still more complex changes [..] "* »

This is the origin of **this small warm little pond**, which has truly made so many researchers in search of the origins of life fantasize. Darwin assumed that his small pond is concentrated enough for the chemistry to progress to the synthesis of a long enough chain of amino acids, the "protein compound".

There may have been small bodies of water on the emerging continents, but were the organic molecules more concentrated there than in the global ocean? Maybe molecules concentrated on the first beaches, or in a few cracks? Were there more organic compounds around the volcanoes? Or at the bottom of the ocean, near hydrothermal vents from which hot gas escapes? Shouldn't we instead imagine particularly effective reactions that make it possible to build polymers (proteins, for example) even under very diluted conditions?

4. The keys to success: energy and catalysis

For a chemical reaction to occur, it is necessary:

that it is possible, it is a matter of thermodynamics,

that it is fast enough, it is a matter of kinetics{ind-text}. in chemistry, kinetics describes the evolution of chemical systems over time, i.e. the passage from an initial state to an end state. The laws of chemical kinetics make it possible to determine the specific velocity of a chemical reaction.{end-tooltip}.

However, a priori, 4 billion years ago, nothing was as it should have been!

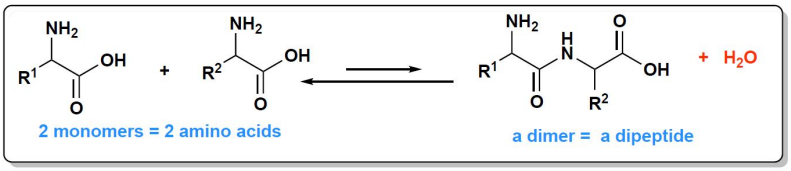


Figure 3. Synthesis of a dipeptide from two amino acids. The difference in the length of the arrows means that the balance is shifted towards the two amino acids. The formation of water, in particular, inhibits the synthesis of dipeptide.

From a thermodynamic point of view, what matters is the relative stability of the starting molecules and the formed molecules. Building a polymer is a step-by-step process. First, two monomers{ind-text}Basic constituents of complex molecules (proteins, complex sugars, nucleic acids, etc.). The successive sequence of these molecules (identical or different), gives rise to a polymer structure. Thus amino acids form proteins, oses form complex sugars, nucleotides form nucleic acids. {end-tooltip} give a dimer, which will be elongated into a trimer and so on, up to very long chains. At first, therefore, two monomers form a dimer and a molecule of water is eliminated. In both peptides (Figure 3) and nucleic acids, dimers are much less stable than monomers. In other words, it is the dimer cut-off reaction (a hydrolysis) that is favoured. The balance is therefore shifted towards the monomers. This especially as this hydrolysis consumes a molecule of water, which in water is favourable, while condensation forms a molecule of water next to the dimer, which is unfavourable. It is this problem of unfavourable formation of a water molecule that leads some authors to seek for the least aqueous environments possible to place their origin of life scenario, the coasts of the first continents in particular, where relatively dry places could undoubtedly be found.

It's no better on the kinetic side. For two molecules to react together, they must be activated, i.e. they must be supplied with a certain amount of energy. The higher the energy to be supplied, the greater is the probability that two monomers will meet (referred to as "shocks") without reacting. In other words, the slower is the reaction. However, the energy required to form an embryonic dimer of protein or nucleic acid is high.

Waste of time? No, since despite all this, it is quite certain that life has appeared. To do this, you need at least:

an energy-rich molecule. By cutting itself into two fragments, this molecule will release a large part of the energy it contains. If this happens simultaneously to the formation of a dimer (e.g. a dipeptide), then the two energies will compensate each other and the overall process will be promoted by thermodynamics;

a catalyst, *i.e.* a chemical species, a molecule, but also sometimes the surface of a solid, (see Origin of the first cells: the engineer's point of view) capable of helping the formation of a dimer from two monomers. The number of effective shocks (those that really form the dimer) will then be much higher, and the reaction will reach a reasonable rate in the prebiotic ocean.

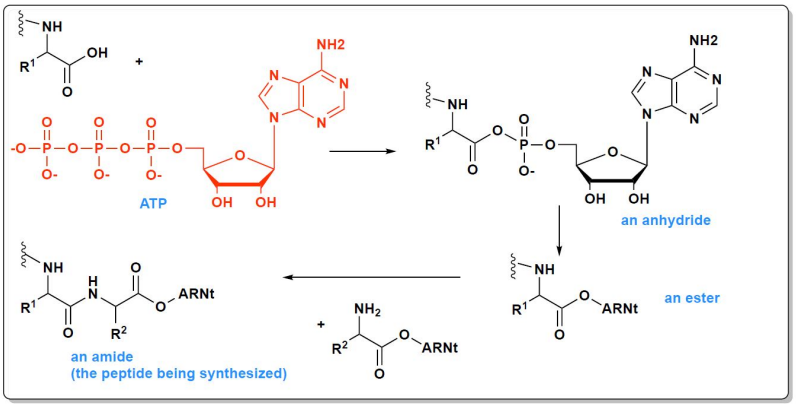


Figure 4. The role of ATP (adenosine triphosphate) in the biological synthesis of peptides The first two reactions are catalysed by a protein, an aminoacyl RNAt synthetase. The third one takes place on the end of two tRNA molecules within the ribosome, which itself is massively composed of RNA chains. That this last stage is catalysed by RNA is a strong argument for the researchers who propose the existence of a primitive world where RNA's (the RNA world) predominated.

The energy-rich molecule used today in living organisms is ATP (Figure 4), **a triphosphate**. The breaking of a phosphate bond releases enough energy to counterbalance the instability of the dimers to be synthesized. In peptide synthesis, this will even allow to reach the dimer *via* intermediates even less stable than the dimer itself. First a mixed phosphoric carboxylic anhydride is formed, then an ester and finally an amide (the peptide). In total, this involves three steps, each of which is inherently slow. Catalysts are therefore involved. Thus, it probably took time for the very first peptides to form in the primitive ocean.

Explaining the appearance of life therefore requires identifying at least one molecular source of energy (knowing that it is highly unlikely that ATP existed in the early ocean, it is too rich in energy), and on the other hand initial catalysts (being understood that those used today in living cells, aminoacyl tRNA synthetases{ind-text}Family of enzymes that catalyse the esterification of amino acids on the 3' end of transfer RNA (tRNA). Preserved in all living organisms, these enzymes help to translate the genetic message into proteins. The amino acids thus added to the end of the tRNAs are then incorporated by the ribosome into the polypeptide chain (protein) being synthesized. There is an aminoacyl tRNA synthetase for each of the 20 amino acids present in proteins. Each of these enzymes recognizes an amino acid and one or more iso-acceptor tRNAs. Their function is essential to the accuracy of the translation of the genetic code, as they ensure that the amino acid thus esterified at the end of the tRNA corresponds to the correct anticodon. {end-tooltip} and others, are much too complex to be prebiotic).

The problem being posed, everything else is only a hypothesis. The most commonly accepted today is that of the "**RNA world**" [3] (see Origin of the first cells: the engineer's point of view). It assumes that the first significant polymers were RNA's, both repositories of first genetic information and catalysts{in-text}Element (organic or mineral) that accelerates or slows down a chemical reaction. Used in very small quantities and specific to a given reaction, the catalyst does not appear in the reaction equation; it does not influence the direction of evolution of the transformation, nor the composition of the system in the final state. An enzyme is a biological catalyst. {end-tooltip}. In fact, some RNA's in current living cells have catalytic properties (although the vast majority of biological catalysts are proteins). With regard to the energy source: it is impossible to form RNA's without using the energy contained in triphosphates. However, it is unlikely that triphosphates could have existed in the early ocean. This is one of the difficulties of the hypothesis. But it has the advantage of reconciling genetic information and catalysis.

On the **other hand, proteins do not carry genetic information, but they are much better catalysts than RNA**. Does this lack of genetic information preclude the idea that proteins could have been the first really important polymers in the history of life? Maybe not... Today indeed, some peptides are manufactured directly on proteins, without the help of nucleic acids. These peptides are called "non-ribosomal" because they are not manufactured in ribosomes{ind-text} A complex composed of RNA and ribosomal proteins, associated with a membrane (at the granular endoplasmic reticulum) or free in the cytoplasm. The function of the ribosome is to translate the genetic code into proteins, through messenger RNAs (mRNAs). The enzymatic activity of the ribosome being carried by rRNAs, the ribosome is a ribozyme. Common to all cells (prokaryotes and eukaryotes), the structure and composition of the ribosome varies according to the organisms. In prokaryotes, the ribosome is said to be 70S (S corresponding to the Sverdberg sedimentation unit) and consists of subunits 50S and 30S. The ribosome of eukaryotes is called 80S, formed by the two subunits 60S and 40S. {end-tooltip}, RNA complexes. However, they are not made "at random" and authors have proposed the existence of a code different from the traditional genetic code (which translates DNA into proteins *via* RNA). This "non-ribosomal" code translates proteins into peptides (one could say peptides into peptides). It is a complex code, based on sets of ten amino acids (in the coding protein): it allows to precisely choose the amino acid to be introduced into the peptide to be synthesized. There is therefore nothing to prevent us from imagining that "pre-genetic" information, even sketchy information, could have been originally carried by amino acid chains.

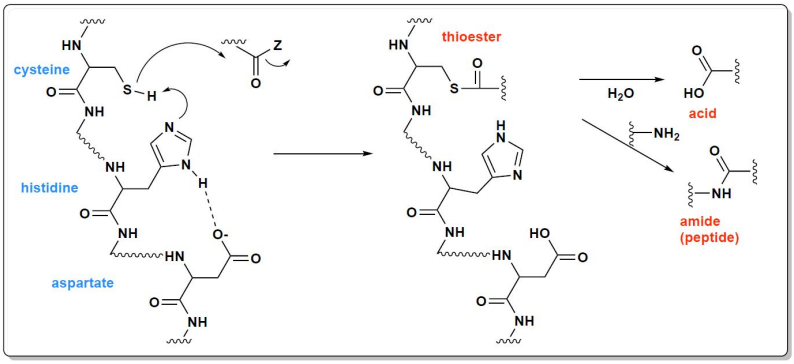


Figure 5. Principle of action of a catalytic triad, here a cysteine - histidine - aspartic acid triad. The zigzag links represent peptide chains. A thioester is formed. It is a high-energy molecule that will be used for subsequent reactions such as hydrolysis. The formation of an amide is another possible follow-up.

Although catalytic proteins are complex molecules, their activity is generally based on fairly simple principles. This is the case for **catalytic triads** found in hydrolases and transferases. Three amino acids are required: an alcohol or z thiol, a base and an acid. In Figure 5, it is a thiol, cysteine, that acts. Thanks to histidine (the base) located further down the protein chain, this cysteine loses its proton. It then carries a negative charge, which allows it to react with the positive carbon of the C=O double bond. Aspartic acid is there to activate histidine. The reaction product is, in this case, **a thioester**, another example, after triphosphates, of a high-energy molecule. This thioester may then undergo other reactions, for example with water to make an acid (the protein that contains the triad is then a hydrolase), or with another organic molecule (the triad is the active site of a transferase).

Of course, in our current proteins, these triads are positioned very precisely by the amino acid chains that bind them. Thanks to this, each triad protein is specialized and only performs one type of reaction on molecules that are also well defined. But, 4 billion years ago, in the primitive ocean, could there be triads? They would undoubtedly have been much less specific, to a certain extent "good at everything". Why not? Figure 6 shows such an example of a very simplified triad. The stereochemistry of the molecule, or "chirality", is specified: it is a very essential question, which a complete model of the origin of life must explain [4].

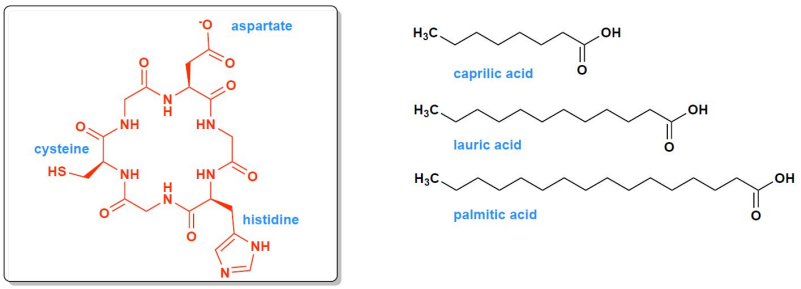


Figure 6. On the left: extreme simplification of a triad. The molecule (hypothetical) is cyclic. It contains the 3 amino acids (cysteine, histidine and aspartic acid). There is no indication that this one would have catalytic activity. On the right: three examples of fatty acids. These are long organic chains, where the CH2 unit repeats itself, terminated by an acid function.

Finally, we cannot imagine life without the existence of partitioned structures, cells or something similar to them, so membranes or walls. The first membranes may have been formed from entangled or more ordered peptides glued together. They may also have contained long hydrophobic organic chains, those of fatty acids (Figure 6). It is possible to synthesize these fatty acids today thanks to proteins and thiolester chemistry.

The world in which life has appeared that is then delineated would be **a world of peptides** much more than a world of nucleic acids. Sulphur, through cysteine and thioesters, would have played a very central role, linking it to a possible world that is even more primitive, more "mineral", **the iron-sulphur world** [5], [6]. This leads us to reflect on the specific role that ferrous iron,

which is soluble, therefore available, could have played in providing electrons, and therefore energy, in relation to all these peptides.

References and notes

[1] https://en.wikipedia.org/wiki/Moon. About the birth of the moon, we will find in this article alternative hypotheses to the giant impact.

[2] Goldford J.E., Hartman H., Smith T.F. & Segré D. (2017) *Remnants of an ancient metabolism without phosphates*. Cell 168, 1-9. http://dx.doi.org/10.1016/j.cell.2017.02.001

[3] https://en.wikipedia.org/wiki/RNA_world

[4] https://en.wikipedia.org/wiki/Chirality_(chemistry)

[5] https://en.wikipedia.org/wiki/Iron-sulfur protein

[6] Valley Y., Shalayel I. et al. (2017) *At the very beginning of life on Earth: The thiol-rich peptide (TRP) world hypothesis.* Int. J. Dev. Biol. 61: 471-478. http://doi.org/10.1387/ijdb.170028yv

L'Encyclopédie de l'environnement est publiée par l'Université Grenoble Alpes - www.univ-grenoble-alpes.fr

Pour citer cet article: **Auteur :** VALLÉE Yannick (2019), Once upon a time when life appeared: chemistry in the earth's ocean 4 billion years ago, Encyclopédie de l'Environnement, [en ligne ISSN 2555-0950] url : http://www.encyclopedie-environnement.org/?p=6844

Les articles de l'Encyclopédie de l'environnement sont mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.