

# At the very beginning of life on Earth: the thiol-rich peptide (TRP) world hypothesis

YANNICK VALLEE<sup>\*,1</sup>, IBRAHIM SHALAYEL<sup>1</sup>, KIEU-DUNG LY<sup>1</sup>, K.V. RAGHAVENDRA RAO<sup>1</sup>, GAEL DE PAËPE<sup>2</sup>,  
KATHARINA MÄRKER<sup>2</sup> and ANNE MILET<sup>1</sup>

<sup>1</sup>Univ. Grenoble Alpes, CNRS, Département de Chimie Moléculaire and

<sup>2</sup>Univ. Grenoble Alpes, INAC, F-38000 Grenoble, France and CEA, INAC, Grenoble, France

**ABSTRACT** Life developed on Earth probably about 3.8 billion years ago, on a planet that was already largely covered by oceans and where the atmosphere was very humid. The reactions, which may have led to the formation of the first polymers, particularly to the first peptides and nucleic acids, must have been compatible with these conditions. This is the case of the reaction of nitriles with aminothiols, such as cysteine and homocysteine. Since aminonitriles are the probable precursors of amino acids, this condensation reaction has been able to rapidly yield dipeptides, tripeptides, oligomers and even true polymers, each containing thiol functions. These thiol-rich peptides (TRP's) would then have assumed the various catalytic roles that the peptides containing cysteine residues play today. They allowed a rapid bloom of life in the primitive ocean. In this scenario, RNA's are not the first polymers, but have been synthesized, like DNA's, thanks to the catalytic properties of thiols in a mostly TRP world. In this world, due to its ability to form a thiolactone, homocysteine may have played the leading role in enabling the previously formed oligomers to be stapled together, thus accelerating the formation of long peptide chains.

**KEY WORDS:** *prebiotic chemistry, thiol rich peptide, metabolism first hypothesis, cysteine, homocysteine*

## Introduction

The answer to the question “what were the first steps of life development on our planet?” largely depends on the answer to this second question: “what were the environmental conditions on Earth in these very old times?” Any attempted answer to the first question incompatible with what is known of these conditions should be excluded. A response only marginally compatible with the general conditions of this environment should be viewed with suspicion.

A second test of the soundness of an origin of life hypothesis should be the speed with which the postulated reactions are expected to lead to catalytic species. Indeed, the faster the implementation of catalytic processes, the more efficient the synthesis of large polymers, from which many will be even more efficient catalysts. Then polymers will interact with polymers. There will be a true bloom of life, with rapid creation of proto-cells (that is, small concentrated monomers and polymers mixes), protecting the formed polymers from too much dispersion and allowing synthetic processes to effectively counterbalance hydrolysis reactions. Saying it in biological terms, anabolism and catabolism together should reach a point of no return.

Furthermore, even though one cannot exclude a completely statistical set of first steps followed by a drastic Darwinian (thermodynamic) molecular selection, retaining a selected collection of polymers after some time, it seems more reasonable to put forward selective processes. The greater the part of actually catalytic oligomers in the initial mixture, the less their activity will be hampered by the presence of unnecessary molecules. This means that the chemistry in action, from monomers to polymers, should better be somehow “informed”. Nowadays information is coded thanks to the genetic code. All nucleic acids are copies of pre-existing nucleic acids. In the synthesis of a given protein, the first Met residue is introduced thanks to the start codon (AUG) and the choice of the second one is coded by another specific codon, let's say AGU for Ser, etc. Nothing is randomly chosen. Of course, in the early days of the development of life, an important part of chance must have existed. However, in this paper we present a scenario that limits the scope of possibilities in this game of chance. It is compatible with initial environmental conditions, it rapidly creates catalytic species and it is selective.

*Abbreviations used in this paper:* TRP, thiol rich peptide.

\*Address correspondence to: Yannick Vallée. Département de Chimie Moléculaire, Univ. Grenoble Alpes, CNRS, Campus Universitaire, 38400 Saint Martin d'Hères, France. E-mail: yannick.vallee@univ-grenoble-alpes.fr -  <http://orcid.org/0000-0002-8412-8989>

Submitted: 9 February, 2017; Accepted: 7 June, 2017.

ISSN: Online 1696-3547, Print 0214-6282

© 2017 UPV/EHU Press (Bilbao, Spain) and Creative Commons CC-BY. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits you to Share (copy and redistribute the material in any medium or format) and Adapt (remix, transform, and build upon the material for any purpose, even commercially), providing you give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. Printed in Spain

## Geological considerations

The story of life probably began *ca.* 3.8 billion years ago, that is just after the late heavy bombardment (Marchi *et al.*, 2014). This bombardment lasted 300 million years, during which thousands of meteorites struck the Earth (Gomes *et al.*, 2005). The energy involved was colossal and it is very unlikely that traces of life, or even complex organic molecules, that would have existed before this event, would have survived it, even though this has been tentatively proposed (Abramov and Mojzsis, 2009, Nutman *et al.*, 2016).

At that time oceans already covered the surface of our planet (Wilde *et al.*, 2001). As the moon was much closer to Earth than nowadays (Walker and Zahnle, 1986), this ocean was shacked by impressive tides. With a Moon-Earth distance 2.8 times shorter than nowadays (Zharkov, 2000), the tides were roughly 22 ( $2.8^3$ ) times stronger. Winds and currents were very strong. It is difficult to have a precise idea of the temperature of this ocean. However, the internal energy of Earth was still much higher than now and probably compensated the lower energy furnished by the Sun, which was smaller (Shaviv, 2003). Thus it could well be that the temperature of water was not too different from nowadays, may be a little bit warmer (Bounama, 2001). As an atmosphere containing a lot of CO<sub>2</sub> covered it, the pH of this ocean might be slightly acidic (Pinti, 2005). The most common ions were Na<sup>+</sup> and Cl<sup>-</sup>, just as it is today. One major difference is that no molecular oxygen, O<sub>2</sub>, was present in the atmosphere, so that the environment was globally reductive (Shaw, 2008). For that reason, iron was in its ferrous state, Fe<sup>2+</sup>, and was soluble. Very active volcanoes (Martin *et al.*, 2007) liberated large amounts of gases, including hydrogen sulfide, H<sub>2</sub>S.

### At the heart of the problem: how to form (catalytic) polymers?

Chemists agree that most, if not all, of the monomeric species necessary to the formation of biological polymers were present (Eschenmoser, 2007), including many amino acids (Bada, 2013, Miller, 1953), nucleic bases (Callahan *et al.*, 2011, Voet and Schwartz, 1983), may be even nucleotides (Becker *et al.*, 2016, Patel *et al.*, 2015). Some may have been synthesized on Earth, over may have been of extraterrestrial origin (Burton *et al.*, 2012), brought by meteorites and asteroids, as well as possibly a part of the water that constituted oceans (Morbidelli *et al.*, 2000). However, after the late heavy bombardment stopped, the arrival of molecules from space inevitably diminished. From that time, the synthesis of most, in not all, prebiotic monomers must have taken place on Earth.

The main problem is then to understand how these monomers condensed to form, *inter alia*, proteins and nucleic acids, a difficult problem indeed when the condensation reactions have to be conducted in water. These reactions are disfavored by thermodynamics (because of the formation of a molecule of water) and kinetics. For instance, the direct condensation in an aqueous solution of an acid with an amine to give an amide is highly improbable (Chalmet *et al.*, 2001).

To circumvent such difficulty, dry reactions have been proposed. We believe that they cannot be taken into account in a realistic prebiotic scenario. One may argue that some reactions occurred on relatively dry surfaces (Cleaves, 2013), on the shore of small continents or near some newly formed streams. But these sur-

faces, if existing, were indeed over wet, due to the large amount of water present in the atmosphere and frequent rains. Many captivating scenarios have been proposed since Darwin's little hot pond (Follmann and Brownson, 2009, Damer, 2016): coastal zone (Lathe, 2003), dehydrating/hydrating cycling (Ross and Damer, 2016), hydrothermal vents (Miller and Bada, 1988), tectonic faults (Schreiber *et al.*, 2012), geothermal fields (Mulkijanian *et al.*, 2012), ocean abysses (Di Giulio, 2005), surface metabolism (Wächtershäuser, 2006), salt-induced peptide formation (Fitz *et al.*, 2007), a stereochemical era (Fontecilla-Camps, 2014), thermal gradients (Agerschou *et al.*, 2017), the sum of various specific spots (Monard, 2016, Stüeken *et al.*, 2013)... To these scenarios, in which life occurred "somewhere", we prefer a scenario in which life came out everywhere or, at least, anywhere in a roughly homogeneous ocean.

We believe that most of organic monomers were randomly distributed in the vigorously stirred global terrestrial ocean and that life began with some efficient reactions in this homogeneous diluted aqueous solution. These reactions had to rapidly yield catalytic oligomers, allowing the synthesis of polymers, then the formation of aggregation micelles, proto-cells, and the first "true" cell models, including the Last Universal Common Ancestor, LUCA (Kyrpides *et al.*, 1999) from which derives the whole tree of life. What could be the key reaction that would lead to the first catalytic oligomers? It has been proposed that the formation of amino acid *N*-carboxyanhydrides (Danger *et al.*, 2012) played a major role in the formation of first peptide chains. However, this process could only explain the formation of a random mixture of peptides. In contrast, our scenario favors the formation of one specific kind of catalytic oligomers, thiol rich peptides (TRP's).

### A plausible answer: the reaction of aminonitriles with aminothiols

Aminonitriles are the first products of the Strecker reaction (Fig. 1A) (Strecker, 1850). These nitriles are subsequently hydrolyzed to amino acids. It is largely admitted that this is the way most amino acids were produced in the early ocean. Nitriles are slightly activated forms of acids and are relatively stable in water. For instance, we found that when aminoacetonitrile, a molecule which has been detected in an interstellar giant molecular cloud (Belloche *et al.*, 2008, Schwell *et al.*, 2012) and was the probable prebiotic precursor of glycine, was dissolved in slightly acidic water at 45°C, it was stable for months (no hydrolysis observed after 3 months). It is thus conceivable that when formed in the primitive ocean, this nitrile and other aminonitriles could react with good nucleophiles before reacting with water.

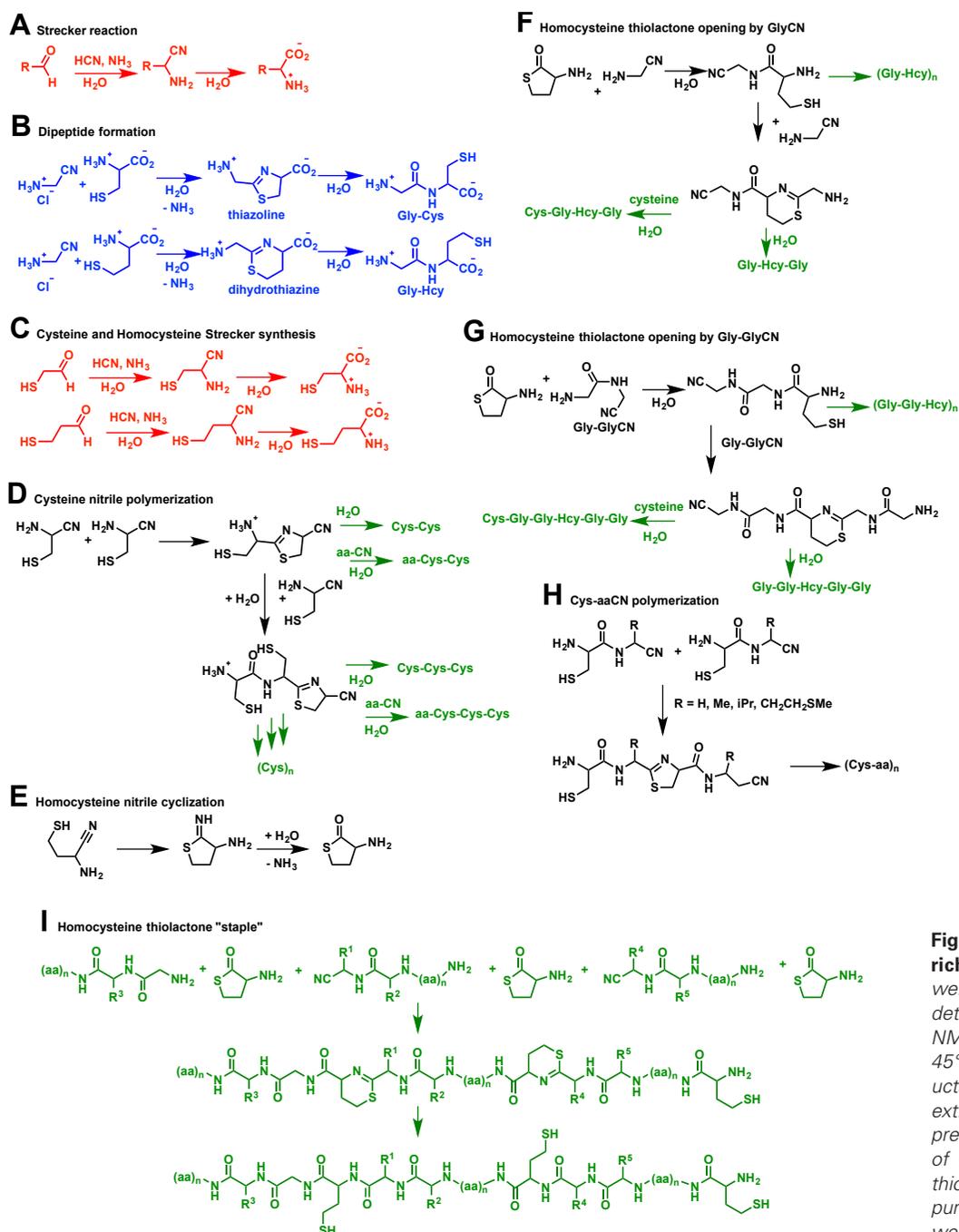
We studied the reaction of aminoacetonitrile with cysteine (Fig. 1B). This reaction first leads to a thiazoline derivative (Krimmer *et al.*, 1987), which is afterward hydrolyzed to yield the dipeptide Gly-Cys. This reaction is very efficient (Fig. 2). It is quicker at 70°C, but goes well also at 25°C. It works at low concentration (we tested it down to 10<sup>-4</sup> mol/L) and at various pH's (at least from pH 4 to 9). Such efficiency is due both to the high nucleophilicity of the thiol group, to the better electrophilicity of aminonitriles compared to other nitriles (a behavior that we confirmed by theoretical calculation) and to the fact that no water molecule is eliminated during the ring formation. Moreover, a water molecule is consumed during the ring opening. It suffers neither insurmountable kinetic obstacle nor

thermodynamic drawback. We found it to work with all aminonitriles (Guillemin and Denis, 1988) we have tested. Homocysteine gave similar results, but no one of the non-thiol amino acids we tested was found to react with nitriles (in water, at 45°C). From these findings, we conclude that aa-Cys and aa-Hcy dipeptides were easily and preferentially formed in the primitive ocean.

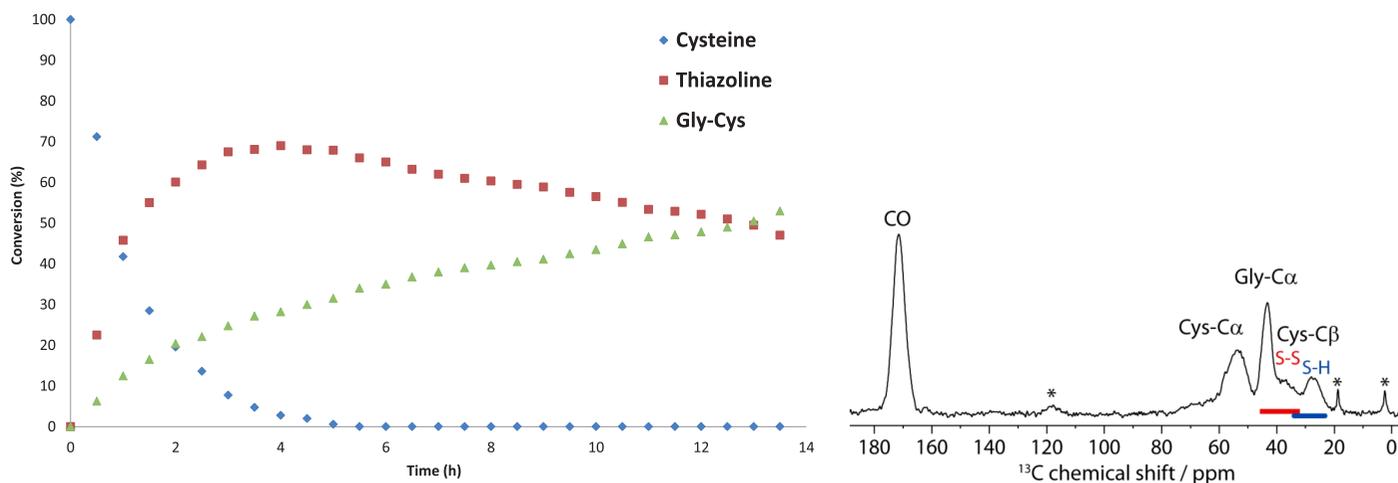
The presence of cysteine and homocysteine in the ocean is highly probable (Parker *et al.*, 2011). As already mentioned, large amounts of H<sub>2</sub>S were delivered during volcanoes eruptions (including under-water eruptions) and were not readily oxidized due to the lack of dioxygen. Mercapto-aldehydes could be formed thanks to the presence of this H<sub>2</sub>S, and these aldehydes led to the

corresponding aminonitriles through the Strecker reaction, hence to cysteine and homocysteine (Fig. 1C). Furthermore, the nitrile precursor of cysteine, which is a plausible interstellar molecule (Knowles *et al.*, 2010), contains both a nitrile group and aminothiol moiety. It is thus expected to be able to polymerize. Indeed, when it was synthesized and kept in water, the formation of Cys-Cys and Cys-Cys-Cys precursors was observed, after what a non-soluble polymer precipitated out of the solution (Fig. 1D).

The behavior of homocysteine nitrile was entirely different (Fig. 1E). When we synthesized it, we did not observe any polymerization, but instead its ring closure to homocysteine thiolactone. This argues for the existence of this thiolactone in the primitive ocean.



**Fig.1 Reactions involved in the thiol-rich peptide scenario.** Experiments were run in H<sub>2</sub>O, for mass spectroscopy determination of products, or in D<sub>2</sub>O, for NMR monitoring. Standard conditions: 45°C, pH slightly acidic (5.5 – 6.5). Products and reactions presented in green are extrapolated from the chemistry used in previous reactions (for instance hydrolysis of thiazoline, opening of homocysteine thiolactone). Used cysteine was enantiopure (L), homocysteine and thiolactone were racemic.



**Fig. 2 (Left).** NMR monitoring of the reaction of cysteine with aminoacetonitrile. The solvent was a 1/1 mixture of  $D_2O$  and  $H_2O$  phosphate buffer (pH 6.5). The NMR probe temperature was  $45^\circ C$ .

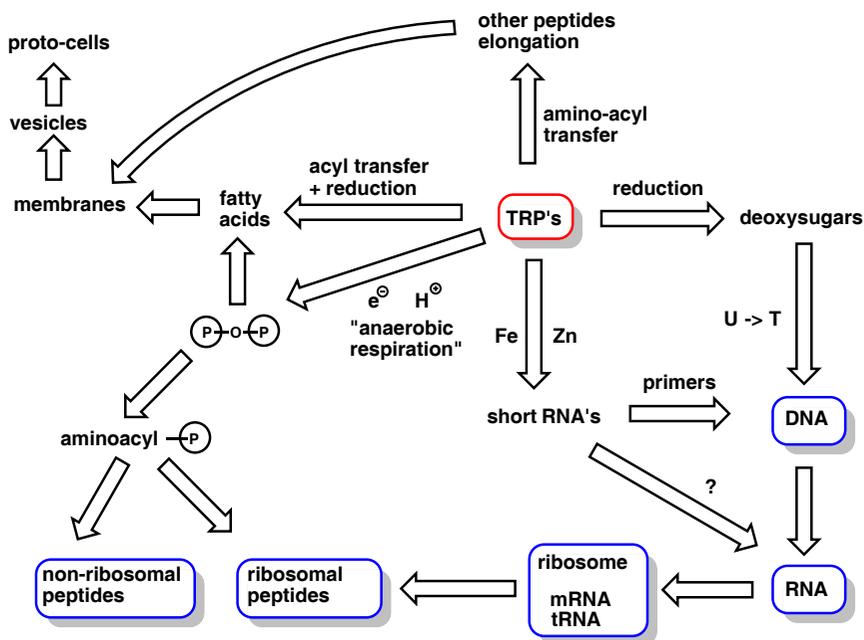
**Fig. 3 (Right).**  $^{13}C$  solid NMR spectrum of poly Cys-Gly. The spectrum was recorded at 9.4 T and a magic-angle spinning frequency of 20 kHz. Red and blue bars denote the expected chemical shift ranges for reduced (S-H) and oxidized (S-S) cysteine residues, indicating the presence of both species in the sample. Peaks labeled with an asterisk correspond to  $CF_3$  (trifluoroacetic acid was used for the synthesis), acetone and silicone grease (from left to right).

Furthermore, when we reacted it with aminoacetonitrile (Fig. 1F), we observed the formation of Gly-Hcy-GlyCN, which is a precursor for Gly-Hcy-Gly, and also for Gly-HCy-Gly-Cys, Gly-HCy-Gly-Hcy and even longer thiol rich peptides (Gly-HCy-Gly-Cys-Cys etc.). Similarly, reacting homocysteine thiolactone with Gly-GlyCN (that we have been able to obtain in a prebiotic manner), we obtained Gly-Gly-HCy-Gly-GlyCN (Fig. 1G). These reactions occurred via the transient formation of Hcy-GlyCN, a precursor for polymeric  $(Hcy-Gly)_n$ , and Hcy-Gly-GlyCN, precursor of  $(Hcy-Gly-Gly)_n$  (Fig. 1G). We are currently studying various Cys-aaCN polymers (Fig. 1H). The  $^{13}C$  solid-state NMR spectrum of poly Cys-Gly obtained this way is presented in Fig. 3. Analysis of the cysteine  $CH_2$  chemical shifts indicates that some cysteine residues have been oxidized to the corresponding disulfides (cysteine) (Sharma and Rajarathnam, 2000).

Overall, we think that these experiments demonstrate that TRP's (thiol-rich peptides) could be readily formed in the prebiotic ocean. Of course, some of their amide bonds could be hydrolyzed, but the hydrolysis of amide bonds is a slow process (the typical half-life of an amide bond in water is ca. 1000 years) (Danger et al., 2012), so that it was efficiently counterbalanced by the specific reactivity of thiol amino acids. At the very most, the hydrolysis vs. amide bond formation game might have favored the most stable oligomers.

Life, as we know it, is inseparable from homochirality. All the models that claim to explain the origin of life must take into account this unavoidable fact. In the case of amino acids, it is possible that some physical factors created a slight excess of the L-form (but it could have been D), for example during partial crystallization (Kojo et al., 2004) or sublimation (Fletcher et al.,

2007). This preference, which was initially very low, may have been sufficient for automultiplication processes to take place (Shibata et al., 1996). Such processes should have been of the utmost importance in the TRP world, which is primarily based on the idea of a quick creation of molecular catalysts. Such catalysts would have been all the more efficient as they would have been homochiral. The choice of the most efficient catalysts, and of the most stable



**Fig. 4.** Proposed scenario from the thiol-rich peptide (TRP) world to the classical DNA/RNA/peptide world. After TRP synthesis, the first steps were the development of a Fe-S based anaerobic respiration, the synthesis of fatty acids, thanks to the acyl carrier properties of TRP's, and the synthesis of short RNAs. DNA's grown from RNA primers when deoxysugars and (may be later) thymine were available. Longer RNAs could be produced either without the support of DNA frameworks or only after the synthesis of DNA strands (and DNA pairing).

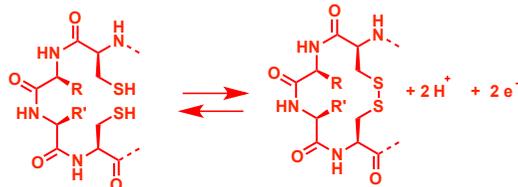
chains, would have favor the synthesis of homochiral peptides. It may be envisaged that this necessitated the implementation of asymmetric Strecker synthesis (Shibasaki *et al.*, 2008).

## Biological considerations

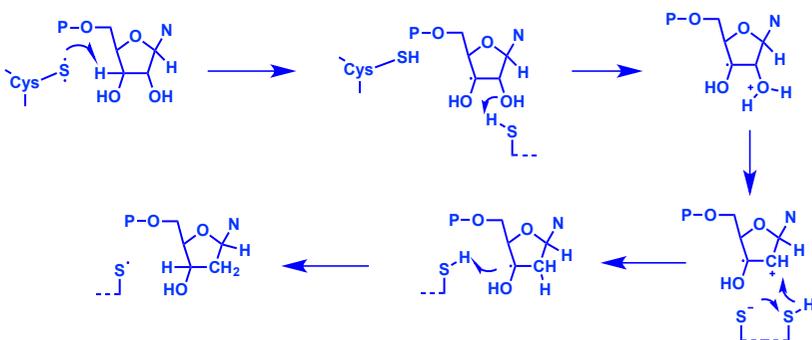
Cysteine rich proteins play an unavoidable role in nowadays biology (Capdevilla and Atrian, 2011, Khalimonchuk and Winge, 2008, Lavergne *et al.*, 2012, Sutherland and Stillman, 2011). We propose that the TRP's presented above assumed at least a part of this role in the primitive ocean and some proportion of the various catalytic functions of cysteine residues in proteins (Fig. 4).

Being able to form disulfides, they took part in redox processes as does, for instance, today's thioredoxin (Holmgren, 1989) (Fig. 5A). As ribonucleotide reductase (Torrents *et al.*, 2002), they were able to reduce sugars, including ribose, reducing it into deoxyribose (Fig. 5B). Being thiols, they were acyl carriers. They were thus able to promote the synthesis of other peptides. Today, non-ribosomal peptides (Hashimoto, 2008) are produced in this way, except that the thiol group of cysteine or homocysteine is replaced by the thiol of a pantetheine group (Fig. 5C) (Keefe *et al.*, 1995, Moffitt and Neilan, 2000). They were also able to built fatty acid chains via malonic condensation of thioesters (Jenke-Kodama, 2005) and reduction of the primary condensation products similarly to today's dehydratase (Buckel *et al.*, 2005). The formation of fatty acids allowed the construction of proto-cell membranes (which probably also contained peptides). Some of these reductions were permitted thanks to the ability of TRP's to form iron clusters (Rawat and Stemmler, 2011) that were able to deliver electrons *via* the  $\text{Fe}^{2+}$  to  $\text{Fe}^{3+}$  oxidation reaction (Fig. 5D). Furthermore, the energy delivered by this iron oxidation might have been used to form polyphosphates (inorganic or organic, similar to ATP). From these polyphosphates ribo and deoxyribonucleotides were formed, as well as acylphosphates, including aminoacyl phosphates, which progressively replaced nitriles in all the previously described reactions. That stage should have been crossed before all HCN present at the beginning of Earth history has been consumed, making the formation of new nitriles impossible. In addition, as the preferentially formed peptides contained a lot of thiol groups, classical native chemical

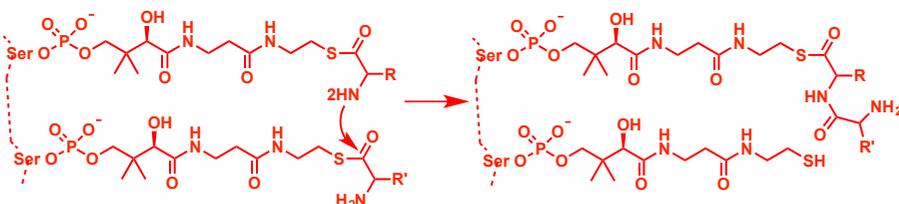
### A Thioredoxine redox chemistry



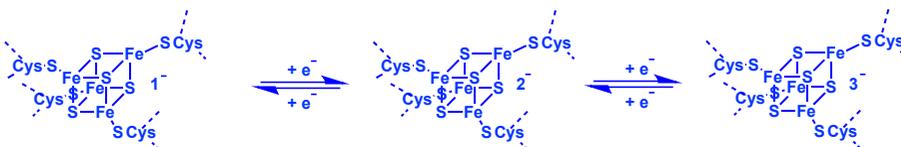
### B Ribonucleotide reductase reduction of riboses into deoxyriboses



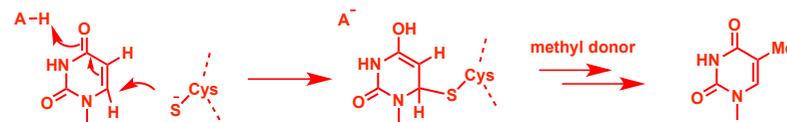
### C Non-ribosomal peptide formation



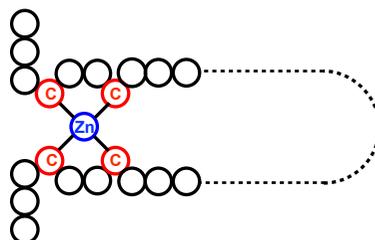
### D Iron clusters redox behavior



### E Uracile to thymine transformation



### F 4-Cys Zinc finger



**Fig. 5. Today's biology that could find some equivalence in the thiol-rich peptide (TRP) world.** Although less efficient than today's proteins, the sulfur based chemistry of TRP's, added to the properties of the other various residues (e.g. Ser, His, Asp...) that were part of their chains, made them the best available catalysts in the early Hadean ocean.

ligation couplings (Dawson *et al.*, 1994) might have taken a part in their elongation.

Also, the nucleophilic thiol of cysteine (or homocysteine) was able to react with the uracil ring to yield the enol, which is mandatory for the methylation of this heterocycle, thus acting as a proto thymidilate synthase (Fig. 5E) (Koehn *et al.*, 2009). The methyl group necessary to complete the transformation of uracil into thymine could come from methylchloride, a compound detected in volcanic gases (Jordan, 2003) or S-methylmethionine, a nowadays common alkylated derivative of methionine (Bourgis *et al.*, 1999). Added to the possible reduction of ribose, this step completes the transformation of RNA sub-units into DNA sub-units. Finally, TRP's were able to promote the ligation of nucleotides. Indeed, cysteine residues, *via* a zinc finger motif (up to six cysteine residues are involved), play a major role in today's nucleic acids – proteins binding (Garvie and Wolberger, 2001), including primase – primer interaction (Griep, 1995). We propose that these zinc fingers (Fig. 5F), as well as iron-sulfur clusters which are also involved (Wu and Brosh, 2012), were formed at a very early stage of life development in a mostly “TRP world” and that they promoted the formation of the first DNA and RNA strands. However, because DNA's are more stable than RNA's (Ross and Deamer, 2016) and can be further stabilized by the formation of double helices, we believe that the first long chain nucleic acids were DNA's, meaning that the first exchange of molecular “genetic” information was the matching of DNA bases (C-G, A-T) during the double helix formation. But ribonucleotides intervened in this polymerization. With long RNA parts, the stability of the DNA-RNA structures was weak and RNA strands were cleaved (second genetic information transmission step, the RNA's thus formed were “genetically” copied from DNA). Some cleaved RNA strands formed protoribosomes and eventually gained some catalytic properties. They interacted with other cleaved strands, proto-mRNA's and proto-tRNA's, two types of molecules themselves interacting *via* codon-anticodon pairing. Aminoacyl-AMP's and peptidyl-AMP's began to react with the termini of proto-tRNA's and a new way to form peptides was discovered. Contrarily to what happened with abiotic TRP's, in which the genetic information (the classical genetic code) was inoperative, but in which some primitive nonribosomal code (Von Döhren, 1999) may have existed, the new synthesis of peptides became completely “informed”.

Because of the specific double reactivity of homocysteine thiolactone (thiolactone opening + aminothiol condensation reaction), Hcy residues may have been more common than Cys residues in TRP's (Jakubowski, 2002). Indeed, Hcy thiolactone can be regarded as a perfect “molecular staple” which would link together two chains of peptidic nitriles (Fig. 1I). The current role of methionine, which is present at the root of any protein synthesis, could be a remembrance of this Hcy dominance, and not, as we proposed before, the legacy of a possible methionine auto-activation (Raghavendra Rao *et al.*, 2015, Vallée *et al.*, 2016). Homocysteine, a nowadays non-proteinogenic amino acid, is still a precursor of methionine. Furthermore, when homocysteine is mistakenly activated by methionyl-tRNA synthase, in place of methionine, it is edited by cyclization into its thiolactone (Jakubowski, 2016), our proposed “staple”.

In the primitive ocean, the formation of a proto-cell first depended on the ability of monomers to grow to polymers, but also on the ability of these polymers to aggregate. Hydrogen bonding

played an important role in this concentration process, as did the hydrophobicity of some amino acid residues. Noticeably, Cys and Hcy residues, which were so numerous in TRP's, are hydrophobic. They are also able to form disulfide bridges. Intramolecular bridges would have stabilized TRP's strands. Intermolecular bridges reinforced the stability of TRP's aggregates. Formation of metallic complexes (*e.g.* with iron) might also have strengthened the structure. The formed aggregates gained greater, more selective, catalytic properties. They became capable of picking monomers from free water, and of promoting their polymerization. Going further into what is of course just a hypothetical scenario, after some time, they even became capable of making monomers. That is to say, in these organic droplets, probably already enclosed in some sort of membrane, a primitive metabolism was running. Finally, when some of these droplets became too big and eventually split into two smaller units, which in turn began to grow, then an archaic form of life was born.

Our proposal is an alternative to nowadays largely consensual RNA world hypothesis (Gilbert, 1986, Liley, 2011, Krupkin *et al.*, 2011). It is a metabolism first hypothesis (Kurland, 2010, Zaia and Zaia, 2008), related to de Duve's thiolester world (De Duve, 2003), and somehow linked to Sutherland's cyanosulfidic protometabolism (Patel *et al.*, 2015), also to Carter's urzyme theory (Li *et al.*, 2013).

## Experimental chemistry

### Reaction of cysteine with aminoacetonitrile

NMR monitored reactions were run in D<sub>2</sub>O solutions, in NMR tubes. For mass measurements, reactions were run in H<sub>2</sub>O. In a typical experiment cysteine (3.9 mg, 0.032 mmol) and aminoacetonitrile hydrochloride (3 mg, 0.032 mmol) were dissolved in D<sub>2</sub>O (750 µl) in a NMR tube under Argon. Starting materials and products were identified in reaction mixtures on the basis of their <sup>1</sup>H and <sup>13</sup>C NMR (500 MHz D<sub>2</sub>O, pH = 6.5). D<sub>2</sub>O calibrated at 4.7 ppm.

Cysteine: <sup>1</sup>H-NMR (ppm): 2.92 – 3.05 (AB of ABX system, 2 H), 4.38 (t, 1 H). <sup>13</sup>C-NMR (ppm): 24.78 (CH<sub>2</sub>-SH), 55.83 (CH), 172.47 (C=O acid). Aminoacetonitrile: <sup>1</sup>H-NMR (ppm): 4.09 (s, 2 H). <sup>13</sup>C-NMR (500 MHz D<sub>2</sub>O) (δ, ppm): 28.06 (CH<sub>2</sub>), 117.27 (CN). Thiazoline: <sup>1</sup>H-NMR (ppm): 3.46 – 3.71 (AB of ABX system, 2 H), 3.83 (s, 2 H), 4.98–5.03 (dd –t, 1 H). <sup>13</sup>C-NMR (ppm): 40.56 (CH<sub>2</sub>-NH<sub>3</sub>), 37.52 (CH<sub>2</sub>-S), 78.97 (CH), 166.84 (C=N Imine), 178.14 (C=O acid). HRMS: calc. 161.0385, found 161.0387. Gly-Cys: <sup>1</sup>H-NMR (ppm): 2.84–2.93 (–d, 2 H), 3.84 (s, 2 H), 4.38 (t, J = 5.3 Hz, 1 H). <sup>13</sup>C-NMR (ppm): 26.09 (CH<sub>2</sub>-NH<sub>3</sub>), 40.51 (CH<sub>2</sub>-SH), 56.83 (CH), 166.6 (C=O amide), 176.06 (C=O acid). HRMS: calc. 179.0490, found 179.0492.

### Reaction of homocysteine thiolactone with aminoacetonitrile

Homocysteine thiolactone (10 mg, 0.065 mmol) and aminoacetonitrile hydrochloride (12 mg, 0.13 mmol) were dissolved in D<sub>2</sub>O (750 µl) in a NMR tube under inert atmosphere, pH = 6. After 22h, the observed conversion into the dihydrothiazine adduct (see Fig. 1F) was 14% (+ 2% Gly-Hcy-GlyCN). After 48h, conversion was 22% (+ 8%). This experiment was repeated in H<sub>2</sub>O for mass spectrometry analysis.

Identification of adduct in the reaction mixture. <sup>1</sup>H-NMR (δ, ppm): 1.56 – 1.62 (m, 1 H of CH<sub>2</sub>), 2.4 – 2.47 (m, 1 H of CH<sub>2</sub>), 3.03 – 3.09 (m, 1 H of CH<sub>2</sub>-S), 3.23 – 3.3 (m, 1 H of CH<sub>2</sub>-S), 4.01 – 4.11 (m, CH), 3.89 (d of AB system, J = 16 Hz, 1 H, CH<sub>2</sub>-NH<sub>3</sub>), 3.82 (d of

AB system,  $J = 16$  Hz, 1 H,  $\text{CH}_2\text{-NH}_3$ ), 4.25 (s,  $\text{CH}_2\text{-CN}$ ).  $^{13}\text{C}$ -NMR ( $\delta$ , ppm): 21.69 ( $\text{CH}_2$ ), 25.17 ( $\text{CH}_2\text{-S}$ ), 27.54 ( $\text{CH}_2\text{-CN}$ ), 44.35 ( $\text{CH}_2\text{-NH}_3$ ), 59.05 (CH), 116.99 (CN), 159.09 (C=N imine), 175.16 (C=O amide). HRMS: calc. 213.0810, found 213.0806.

A full account of experimental details will be presented in a forthcoming chemistry paper.

#### Author Contributions

Y.V. conceived the TRP world theory and supervised the experimental work. I.S., L.K.D. and K.V.R.R. performed the experiments. G.d.P. and K.M. performed and interpreted the solid-state NMR experiments. A.M. was in charge of theoretical calculations.

#### Acknowledgments

We thank le Ministère des Affaires Étrangères (Consulat Général de France à Jérusalem) for IS grant, la Mairie de Gières for their help and the Laboratoire d'Excellence ARCANÉ (ANR-11-LABX-0003-01) for funding through KVVR grant.

#### References

- ABRAMOV, O. and MOJSIS, S.J. (2009). Microbial habitability of the Hadean Earth during the late heavy bombardment. *Nature* 459: 419-422.
- AGERSCHOU, E.D., MAST, C.B. and BRAUN, D. (2017). Emergence of life from trapped nucleotides? Non-equilibrium behavior of oligonucleotides in thermal gradients. *Synlett*. 27: 56-63.
- BADA, J. L. (2013). New insights into prebiotic chemistry from Stanley Miller's spark discharge experiments. *Chem. Soc. Rev.* 42: 2186-2196.
- BECKER, S., THOMA, I., DEUTSCH, A., GEHRKE, T., MAYER, P., ZIPSE, H. and CARELL, T. (2016). A high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. *Science* 352: 833-836.
- BELLOCHE, A., MENTEN, K.M., COMITO, C., MÜLLER, H.S.P., SCHILKE, P., OTT, J., THORWIRTH, S. and HIERET, C. (2008). Detection of amino acetonitrile in Sgr B2(N). *Astron. Astrophys.* 482: 179-196.
- BOUNAMA, C., FRANCK, S. and VON BLOH, W. (2001). The fate of Earth's ocean. *Hydrol. Earth System. Sc.* 5: 569-575.
- BOURGIS, F., ROJE, S., NUCCIO, M.L., FISCHER, D.B., TARCZYNSKI, M.C., LI, C., HERSCHBACH, C., RENNENBERG, H., PIMENTA, M. J., SHEN, T.L., GAGE, D.A. and HANSON, A. D. (1999). S-Methylmethionine plays a major role in phloem sulfur transport and is synthesized by a novel type of methyltransferase. *Plant Cell* 11: 1485-1497.
- BUCKEL, W., MARTINS, B.M., MESSERSCHMIDT, A. and GOLDING, B.T. (2005). Radical mediated dehydration reactions in anaerobic bacteria. *Biol. Chem.* 386: 951-959.
- BURTON, A.S., STERN, J.C., ELSIBA, J.E., GLAVIN, D.P. and DWORKIN, J.P. (2012). Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites. *Chem. Soc. Rev.* 41: 5459-5472.
- CALLAHAN, M.P., SMITH, K.E., CLEAVES II, H.J., RUZICKA, J., STERN, J.C., GLAVIN, D.P., HOUSE, C.H. and DWORKIN, J.P. (2011). Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. *Proc. Natl. Acad. Sci. USA* 108: 13995-13998.
- CAPDEVILLA, M. and ATRIAN, S. (2011). Metallothionein protein evolution: a miniasay. *J. Biol. Inorg. Chem.* 16: 977-989.
- CHALMET, S., HARB, W. and RUIZ-LÓPEZ, M.F. (2001). Computer simulation of amide bond formation in aqueous solution. *J. Phys. Chem. A* 105: 11574-11581.
- CLEAVES II, H.J.C. (2013). Prebiotic chemistry: geochemical context and reaction screening. *Life* 3: 331-345.
- DAMER, B. (2016). A field trip to the Archean in search of Darwin's warm little pond. *Life* 6: 21.
- DANGER, G., PLASSON, R. and PASCAL, R. (2012). Pathways for the formation and evolution of peptides in prebiotic environments. *Chem. Soc. Rev.* 41: 5416-5429.
- DAWSON, P.E., MUIR, T.W., CLARK-LEWIS, I. and KENT, S.B. (1994). Synthesis of proteins by native chemical ligation. *Science* 266: 776-778.
- DE DUVE, C. (2003). A research proposal on the origin of life. *Orig. Life Evol. Biosph.* 33: 559-574.
- DI GIULIO, M. (2005). The ocean abysses witnessed the origin of the genetic code. *Gene* 346: 7-12.
- ESCHENMOSER, A. (2007). The search for the chemistry of life's origin. *Tetrahedron* 63: 12821-12844.
- FITZ, D., REINER, H. and RODE B.M. (2007). Chemical evolution toward the origin of life. *Pure Appl. Chem.* 79: 2101-2117.
- FLETCHER, S.P., JAGT, R.B.C. and FERINGA, B.L. (2007). An astrophysically-relevant mechanism for amino acid enantiomer enrichment. *Chem. Commun.* 2578-2580.
- FOLLMANN, H. and BROWNSON, C. (2009). Darwin's warm little pond revisited: from molecules to the origin of life. *Naturwissenschaften* 96: 1265-1292.
- FONTECILLA-CAMPS, J.C. (2014). The stereochemical basis of the genetic code and the (mostly) autotrophic origin of life. *Life* 4, 1013-1025.
- GARVIE, C.W. and WOLBERGER, C. (2001). Recognition of specific DNA sequences. *Mol. Cell* 8: 937-946.
- GILBERT, W. (1986). The RNA world. *Nature* 319: 618.
- GOMES, R., LEVISON, H.F., TSIGANIS, K. and MORBIDELLI, A. (2005). Origin of the cataclysmic late heavy bombardment period of the terrestrial planets. *Nature* 435: 466-469.
- GRIEP, M.A. (1995). Primase structure and function. *Ind. J. Biochem. Biophys.* 32: 171-178.
- GUILLEMIN, J.C. and DENIS, J.M. (1988). Synthèse d'imines linéaires non-stabilisées par réactions gaz-solide sous vide. *Tetrahedron* 44: 4431-4446.
- HASHIMOTO, S.-I. (2008). Ribosome-independent peptide synthesis and their application to peptide production. *J. Biol. Macromol.* 8: 28-37.
- HOLMGREN, A. (1989). Thioredoxin and glutaredoxin systems. *J. Biol. Chem.* 264: 13963-13966.
- JAKUBOWSKI, H. (2002). Homocysteine is a protein amino acid in humans. *J. Biol. Chem.* 277: 30425-30428.
- JAKUBOWSKI, H. (2016). Aminoacyl-tRNA synthetases and the evolution of coded peptide synthesis: the thioester world. *FEBS Lett.* 590: 469-481.
- JENKE-KODAMA, H., SANDMANN, A., MÜLLER, R. and DITTMANN, E. (2005). Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.* 22: 2027-2039.
- JORDAN, A. (2003). Volcanic formation of halogenated organic compounds. In *The handbook of environmental chemistry*, Vol. 3P (Ed. G.W.Gribble). Springer-Verlag, Berlin, pp.121-139.
- KHALIMONCHUK, O. and WINGE, R. (2008). Function and redox state of mitochondrial localized cysteine-rich proteins important in the assembly of cytochrome c oxidase. *Biochim. Biophys. Acta* 1783: 618-628.
- KEEFE, A.D., NEWTON, G.L. and MILLER, S.L. (1995). A possible prebiotic synthesis of pantetheine, a precursor to coenzyme A. *Nature* 373: 683-685.
- KNOWLES, D.J., WANG, T. and BOWIE, J.H. (2010). Radical formation of amino acid precursors in interstellar regions? Ser, Cys and Asp. *Org. Biomol. Chem.* 8: 4934-4939.
- KOEHN, E.M., FLEISCHMANN, T., CONRAD, B.A., LESLEY, S.A. MATHEWS, I.I. and KOHEN, A. (2009). An unusual mechanism of thymidylate biosynthesis in organisms containing the thyX gene. *Nature* 458: 919-923.
- KOJO, S., UCHINO, H., YOSHIMURA, M. and TANAKA, K. (2004). Racemic D,L-asparagine causes enantiomeric excess of other coexisting racemic D,L-amino acids during recrystallization: a hypothesis accounting for the origin of L-amino acids in the biosphere. *Chem. Commun.* 2146-2147.
- KRIMMER, H.P., DRAUZ, K. and KLEEMANN, A. (1987). Umsetzung von  $\beta$ -mercapto- $\alpha$ -aminosäuren mit nitrilen. *Chemiker-Zeitung* 111: 357-361.
- KRUPKIN, M., MATZOV, D., TANG, H., METZ, M., KALAORA, R., BELOUSOFF, M.J., ZIMMERMAN, E., BASHAN, A. and YONATH, A. (2011). A vestige of a prebiotic bonding machine is functioning within the contemporary ribosome. *Phil. Trans. R. Soc. B* 366: 2972-2978.
- KURLAND, C.G. (2010). The RNA dreamtime. *Bioessays* 3: 866-871.
- KYRPIDES, N., OVERBEEK, R. and OUZOUNIS, C. (1999). Universal protein families and the functional content of the last universal common ancestor. *J. Mol. Evol.* 49: 413-423.
- LATHE, R. (2004). Fast tidal cycling and the origin of life. *Icarus* 168: 18-22.
- LAVERGNE, V., TAFT, R. J. and ALEWOOD, P.F. (2012). Cystein-rich mini-proteins

- in human biology. *Cur. Top. Med. Chem.* 12: 1514-1533.
- LI, L., FRANCKLYN, C. and CARTER, C.W. (2013). Aminoacylating urzymes challenge the RNA world hypothesis. *J. Biol. Chem.* 288: 26856-26863.
- LILLEY, D.M.J. (2011). Mechanism of RNA catalysis. *Phil. Trans. R. Soc. B* 366: 2910-2917.
- MARCHI, S., BOTTKE, W.F., ELKINS-TANTON, L.T., BIERHAUS, M., WUENNEMANN, K., MORBIDELLI, A. and KING, D.A. (2014). Widespread mixing and burial of Earth's Hadean crust by asteroid impacts. *Nature* 511: 578-582.
- MARTIN, R.S., MATHER, T.A. and PYLE, D.M. (2007). Volcanic emissions and the early Earth atmosphere. *Geochim. Cosmochim. Ac.* 71: 3673-3685.
- MILLER, S.L. (1953). A production of amino acids under possible primitive Earth conditions. *Science* 117: 528-529.
- MILLER, S.L. and BADA, J.L. (1988). Submarine hot springs and the origin of life. *Nature* 334: 609-611.
- MOFFITT, M.C. and NEILAN, B.A. (2000). The expansion of mechanistic and organismic diversity associated with non-ribosomal peptides. *FEMS Microbiol. Lett.* 191: 159-167.
- MONNARD, P.A. (2016). Taming prebiotic chemistry: The role of heterogeneous and interfacial catalysis in the emergence of a prebiotic catalytic/information polymer system. *Life* 6: 40.
- MORBIDELLI, A., CHAMBERS, J., LUNINE, J.I., PETIT, J.M., ROBERT, F., VALSECHI, G.B. and CYR, K.E. (2000). Source regions and timescales for the delivery of water to the Earth. *Meteorit. Planet. Sci.* 35: 1309-1320.
- MULKIDJANIAN, A.Y., BYCHKOV, A.Y., DIBROVA, D.V., GALPERIN, M.Y. and KOONIN, E.V. (2012). Origin of first cells at terrestrial, anoxic geothermal fields. *Proc. Natl. Acad. Sci. USA* 109: E821-E830.
- NUTMAN, A.P., BENNETT, V.C., FRIEND, C.R.L., VAN KRANENDONK, M.J. and CHIVAS, A.R. (2016). Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature* 537: 535-538.
- PATEL, B.H., PERCIVALLE, C., RITSON, D.J., DUFFY, C.D. and SUTHERLAND, J.D. (2015). Common Origin of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nature Chem.* 7: 301-307.
- PARKER, E.T., CLEAVES, H.J., DWORKIN, J.P., GLAVIN, D.P., CALLAHAN, M., AUBREY, A., LAZCANO, A. and BADA, J.L. (2011). Primordial synthesis of amines and amino acids in a 1958 Miller H<sub>2</sub>S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. USA* 108: 5526-5531.
- PINTI, D.L. (2005). The origin and evolution of the oceans. In *Lectures in astrobiology*, Vol. 1 (Eds M Gargaud, B Barbier, H Martin and J Reisse). Springer-Verlag, Berlin Heidelberg, pp. 83-112.
- RAGHAVENDRA RAO, K.V., CAIVEAU, N., DAVID, R., SHALAYEL, I., MILET, A. and VALLÉE Y. (2015). Theoretical study, synthesis, and reactivity of five-membered acyl sulfonium cations. *Eur. J. Org. Chem.* 6125-6129.
- RAWAT, S. and STEMMLER, T.L. (2011). Key players and their role during mitochondrial Iron-Sulfur cluster biosynthesis. *Chem. Eur. J.* 17: 746-753.
- ROSS, D.S. and DEAMER, D. (2016). Dry/wet cycling and the thermodynamics and kinetics of prebiotic polymer synthesis. *Life* 6: 26; DOI: 10.3390/life6030028.
- SAAVIV, N.J. (2003). Towards a solution to the early faint Sun paradox: a lower cosmic ray flux from a stronger solar wind. *J. Geophys. Res.* 108: 1437.
- SHARMA, D. and RAJARATHNAM, K. (2000). <sup>13</sup>C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* 18: 165-171.
- SHAW, H.G. (2008). Earth's atmosphere – Hadean to early proterozoic. *Chemie der Erde* 68: 235-264.
- SCHREIBER, U., LOCKER-GRÜTJEN, O. and MAYER C. (2012). Hypothesis: origin of life in deep-reaching faults. *Orig. Life Evol. Biosph.* 42: 47-54.
- SCHWELL, M., GAIE-LEVREL, F., BÉLINAN, Y., GAZEAU, M.C., FRAY, N., SAUL, G., CHAMPION, N., LEACH, S. and GUILLEMIN, J.C. (2012). VUV spectroscopy and photochemistry of five interstellar and putative prebiotic molecules. In *European conference and laboratory astrophysics*, ECLA, EAS publication series 58 (Eds C Stehél, C Joblin and L d'Hendecourt). EDP Sciences, pp. 301-306.
- SHIBATA, T., MORIOKA, H., HAYASE, T., CHOJI, K. and SOAI, K. (1996). Highly enantioselective catalytic asymmetric automultiplication of chiral pyrimidyl alcohol. *J. Am. Chem. Soc.* 118: 471-472.
- SHIBASAKI, M., KANAI, L. and Mita, T. (2008). The catalytic asymmetric Strecker reaction. *Org. Reactions* 70: 1-119.
- STRECKER, A. (1850). Ueber die künstliche bildung der milchsäure und einen neuen, dem glycocoll homologen körper. *Liebigs Ann. Chem.* 75: 27-45.
- STÜEKEN, E.E., ANDERSON, R.E., BOWMAN, J.S., BRAZELTON, W.J., COLANGELO-LILLIS, J., GOLDMAN, A.D., SOM, S.M. and BAROSS, J.A. (2013). Did life originate from a global chemical reactor? *Geobiology* 11: 101-126.
- SUTHERLAND, D.E.K. and STILLMAN, M.J. (2011). The "magic numbers" of metallothionein. *Metallomics* 3: 444-463.
- TORRENTS, E., ALOY, P., GILBERT, I. and RODRIGUEZ-TRELLES, F. (2002). Ribonucleotide reductases: divergent evolution of an ancient enzyme. *J. Mol. Evol.* 55: 138-152.
- VALLÉE, Y., MILET, A. and RAGHAVENDRA RAO, K.V. (2016). Was methionine the molecular trigger of life on Earth? *Phosphorus Sulfur*. 191: 329-332.
- VOET, A.B. and SCHWARTZ, A.W. (1983). Prebiotic adenine synthesis from HCN – Evidence for a newly discovered major pathway. *Bioorg. Chem.* 12: 8-17.
- VON DÖHREN, H., DIECKMANN, R. and PAVELA-VRANCIC, M. (1999). The non-ribosomal code. *Chem. Biol.* 6: R273-279
- WÄCHTERSCHÄUSER, G. (2006). From volcanic origins of chemoautotrophic life to Bacteria, Archaea and Eukarya. *Phil. Trans. R. Soc. B* 361: 1787-1808.
- WALKER, J.C.G. and ZAHNLE, K.J. (1986). Lunar nodal tide and distance to the Moon during the Precambrian. *Nature* 320: 600-602.
- WILDE, S.A., VALLEY, J.W., PECK, W.H. and GRAHAM, C.M. (2001). Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature* 409: 175-178.
- WU, Y. and BROSH, R.M. Jr (2012). DNA helicase and helicase-nuclease enzymes with a conserved iron-sulfur cluster. *Nucleic Acids Res.* 40: 4247-4260.
- ZAIA, D.A.M. and ZAIA, C.T.B.V. (2008). Algumas controvérsias sobre a origem da vida. *Quim. Nova* 31: 1599-1602.
- ZHARKOV, V.N. (2000). On the history of the lunar orbit. *Solar System Res.* 34: 1-11.

**Further Related Reading, published previously in the *Int. J. Dev. Biol.***

**Comparative expression analysis of cysteine-rich intestinal protein family members *crip1*, *2* and *3* during *Xenopus laevis* embryogenesis**

Annemarie Hempel and Susanne J. Kühl  
Int. J. Dev. Biol. (2014) 58: 841-849  
<https://doi.org/10.1387/ijdb.140270sk>

**N-acetylcysteine protects induced pluripotent stem cells from in vitro stress: impact on differentiation outcome**

Ina Berniakovich, Leopoldo Laricchia-Robbio and Juan Carlos Izpisua Belmonte  
Int. J. Dev. Biol. (2012) 56: 729-735  
<https://doi.org/10.1387/ijdb.120070ji>

**Participation of cysteine-rich secretory proteins (CRISP) in mammalian sperm-egg interaction**

Débora J. Cohen, Dolores Busso, Vanina Da Ros, Diego A. Ellerman, Julieta A. Maldera, Nadia Goldweic and Patricia S. Cuasnicu  
Int. J. Dev. Biol. (2008) 52: 737-742  
<https://doi.org/10.1387/ijdb.072538dc>

**A hypothesis linking low folate intake to neural tube defects due to failure of post-translation methylations of the cytoskeleton**

Natalie K. Björklund and Richard Gordon  
Int. J. Dev. Biol. (2006) 50: 135-141  
<https://doi.org/10.1387/ijdb.052102nb>

**Cysteine-rich region of X-Serrate-1 is required for activation of Notch signaling in *Xenopus* primary neurogenesis.**

Tomomi Kiyota and Tsutomu Kinoshita  
Int. J. Dev. Biol. (2002) 46: 1057-1060  
<http://www.intjdevbiol.com/web/paper/12533029>

**Crescent, a novel chick gene encoding a Frizzled-like cysteine-rich domain, is expressed in anterior regions during early embryogenesis.**

P L Pfeffer, E M De Robertis and J C Izpisua-Belmonte  
Int. J. Dev. Biol. (1997) 41: 449-458  
<http://www.intjdevbiol.com/web/paper/9240561>

**5 yr ISI Impact Factor (2013) = 2.879**

