

Article

Did Homocysteine Take Part in the Start of the Synthesis of Peptides on the Early Earth?

 Sparta Youssef-Saliba , Anne Milet  and Yannick Vallée * 

University Grenoble Alpes, CNRS, DCM, F-38058 Grenoble, France; sparta_saliba@hotmail.com (S.Y.-S.); anne.milet@univ-grenoble-alpes.fr (A.M.)

* Correspondence: yannick.vallee@univ-grenoble-alpes.fr

Abstract: Unlike its shorter analog, cysteine, and its methylated derivative, methionine, homocysteine is not today a proteinogenic amino acid. However, this thiol containing amino acid is capable of forming an activated species intramolecularly. Its thiolactone could have made it an interesting molecular building block at the origin of life on Earth. Here we study the cyclization of homocysteine in water and show theoretically and experimentally that in an acidic medium the proportion of thiolactone is significant. This thiolactone easily reacts with amino acids to form dipeptides. We envision that these reactions may help interpret why a methionine residue is introduced at the start of all protein synthesis.

Keywords: prebiotic chemistry; abiogenesis; homocysteine; thiolactone; peptides



Citation: Youssef-Saliba, S.; Milet, A.; Vallée, Y. Did Homocysteine Take Part in the Start of the Synthesis of Peptides on the Early Earth?.

Biomolecules **2022**, *12*, 555. <https://doi.org/10.3390/biom12040555>

Academic Editors: Anton Hermann and Guzel F. Sitdikova

Received: 22 March 2022

Accepted: 6 April 2022

Published: 8 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

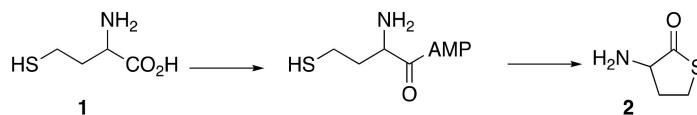


Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

How life came to be on Earth remains a mystery. However, the large number of relevant results published to date allow for some plausible hypotheses [1]. As the first steps that would eventually lead to life were taking place, that is ca. four billion years ago, the main source of carbon on our planet was carbon dioxide [2]. It is thus reasonable to assume that most (if not all) of today's organic carbon comes from this initial CO₂ [3]. However, other sources of carbon have been proposed, particularly cyanides [4]. Furthermore, it is possible that some of the earliest bricks of life were not made on Earth but delivered from space at a time when our planet was still undergoing massive meteorite bombardment [5].

Although it does not participate in building proteins, homocysteine **1** is an important amino acid in the current metabolism. Its biochemistry is intimately linked to those of the two proteionogenic sulfur amino acids, cysteine and methionine [6]. When it is mistakenly introduced into the process of synthesizing a peptide in the ribosome, it is rejected by an editing mechanism that produces its thiolactone **2** (Scheme 1) [7].



Scheme 1. Formation of thiolactone **2** from **1**.

This editing is important since it avoids the synthesis of proteins which would have abnormal behaviors, potentially leading to dysfunctions and to serious diseases [8]. However, in absolute terms, nothing prevents imagining proteins in which all or part of the Cys residues would have been replaced by Hcy residues. The thiol function of homocysteine might play a role similar to that of its shorter counterpart. Hcy residues could form disulfide bridges, enter in the composition of metal complexes and lead to some radical chemistry. However, although plausible prebiotic syntheses of this amino acid have been

reported [9,10], for some reason, probably at an early stage of its development, life rejected the possibility to introduce it in proteins.

If, due to the specificity of their thiol function, Cys residues play a considerable role in today's proteins, especially in multiple catalytic processes, Met residues can generally be considered as lipophilic residues little different from others (Leu, Ile, Val...) and easily interchangeable with them, although sometimes their sulfur atom chelates a metal [11]. No doubt, the most striking role of Met is that it starts the synthesis of all archaea and eukaryotes proteins, and that its *N*-formylated derivative starts the synthesis of all bacteria proteins. It has been proposed that this choice, if not performed purely at random, could be correlated with the relative rarity of the Met residue (2.32% in a representative set of proteins [12]) and with the fact that it is encoded only by a single codon (AUG). Compared to a more common residue and/or a residue encoded by several codons, this would be a better indication to signify a start residue. A notable difficulty here is that the consensus on the order of introduction of the 20 usual residues in the genetic code ranks methionine as one of the very last amino acids to enter the code [13]. There would therefore have been a period with another start residue or without a defined start residue.

In our genetic world, the real starting point is not so much Met but its codon, AUG. However, many exceptions to the use of AUG are known. For example, 17% of the start codons in *Escherichia coli* are different from AUG (mostly GUG, UUG) [14,15]. However, independent of the used start codon, fMet always starts the synthesis of peptides. That is, fMet looks more invariable than AUG. This variability of the start codon may be the result of late adaptations in the history of life. Alternatively, it could mean that Met (or fMet) was chosen before the start codon, which then should have adapted to the presence of the start Met. However, nothing in the chemistry of methionine predisposes it to be more reactive than other amino acids. This is not the case with its demethylated analog, homocysteine, which can form its thiolactone. Such an activated species could have played an important role in the thioester world proposed by de Duve [16].

In this article, we study the cyclization in water of homocysteine into its thiolactone, first experimentally, then using theoretical calculations. We also present our results about the reaction of this thiolactone and of two of its *N*-acylated derivatives with amino acids to form dipeptides.

2. Materials and Methods

¹H-NMR and ¹³C-NMR spectra were recorded on a Bruker Avance 500 (Billerica, MA, USA) (1H:500 MHz, 13C:125 MHz) and 400 (1H:400 MHz, 13C:100 MHz) spectrometers.

High-resolution mass spectra (HRMS) were recorded on LTQ Orbitrap XL (Thermo Scientific, Illkirch, France) mass spectrometer.

Homocysteine **1**, thiolactone **2** (as its HCl salt), *N*-acetylthiolactone **6** and all used amino acids are commercially available. They were purchased and used without further purification.

Synthesis of *N*-formyl thiolactone **4** [17]. Acetic anhydride (60 mL) was mixed with 30 mL of formic acid under inert atmosphere. The solution was stirred at room temperature for 24 h. A total of 10 g of homocysteine thiolactone was added and the reaction left to stir under reflux for 5 h. An oil was obtained after vacuum elimination of the solvents. This oil was purified by flash column chromatography (90:10 ethyl acetate:pentane); **4** was obtained in a 68% yield.

¹H-NMR (500 MHz, CDCl₃): 8.19 (s, 1H, C(O)H), 6.57 (s, 1H, NH), 4.55 (m, 1H), 3.27 (m, 2H, CH₂), 2.85 (m, 1H of CH₂), 1.93 (m, 1H of CH₂). ¹³C-NMR (125 MHz, CDCl₃): 205.2, 161.8, 58.1, 31.7, 27.6. HRMS (ESI): Calcd for C₅H₈O₂NS [M + H]⁺: *m/z* = 146.02703, found: *m/z* = 146.02669.

Most reactions were run in NMR tubes with no purification of products. The obtained products were identified in reaction mixtures. Only in the case of glycine, preparative reactions were run which permitted the isolation of formed dipeptides.

2 + glycine. Homocysteine thiolactone **2** HCl salt (5 mg, 0.033 mmol) was mixed with 2 equivalent of glycine (4.9 mg, 0.066 mmol) in 5 mL of degazed distilled water. The solution was basified to pH 8 by adding NaHCO₃. The mixture was stirred at 45 °C for

24 h. The crude product was purified on an inverse phase silica preparative column (50:50:1 H₂O:ACN:TFA). Hcy-Gly was isolated in a 26% yield.

¹H-NMR (500 MHz, D₂O): 4.07 (dd, *J* = 4, 7 Hz, 1H), 3.75 (dd of AB system, *J* = 17, 30 Hz, 2H), 2.82–2.71 (m, 2H), 2.28–2.23 (m, 2H). ¹³C-NMR (125 MHz, D₂O): 175.99, 168.98, 52.52, 43.39, 31.80, 29.99. HRMS (ESI): Calcd for C₁₂H₂₁N₄O₆S₂ (disulfide) [M–H][−]: *m/z* = 381.09080, found: *m/z* = 381.09088.

4 + glycine. Similarly, we isolated *N*-formyl-Hcy-Gly, from the reaction of **4** with glycine (yield 56%).

¹H-NMR (500 MHz, D₂O): 8.10 (s, 1H, C(O)H), 4.56 (dd, *J* = 6, 9 Hz, 1H), 3.68 (m, 2H), 2.54 (m, 2H, CH₂), 1.98 (m, 2H, CH₂). ¹³C-NMR (125 MHz, D₂O): 177.3, 161.2, 59.2, 40.5, 32.4, 28.2. HRMS (ESI): Calcd for C₇H₁₁N₂O₄S [M–H][−]: *m/z* = 219.04450, found: *m/z* = 219.04460.

6 + glycine. *N*-acetyl-Hcy-Gly was isolated in a 46% yield. ¹H-NMR (500 MHz, D₂O): 4.42 (dd, *J* = 6, 9 Hz, 1H), 3.92 (m, 2H), 2.82–2.65 (m, 2H), 2.21–1.97 (m, 2H), 1.95 (s, 3H). ¹³C-NMR (125 MHz, D₂O): 174.36, 174.12, 172.93, 52.36, 41.04, 33.74, 30.46, 21.71. HRMS (ESI): Calcd for C₁₆H₂₅N₄O₈S₂ (disulfide) [M–H][−]: *m/z* = 465.11193, found: *m/z* = 465.11168.

Dimerization of **2** on silica. Homocysteine **1** (5 mg) was diluted in a minimum amount of water (less than 0.1 mL). An excess of silica was added to the solution. Water was then evaporated under vacuum with heating at 60 °C. The obtained dry solid was then left overnight at 80 °C. D₂O was then added, and the mixture was filtered. The D₂O phase was collected and analyzed by ¹H NMR. It was consistent with a mixture 1/2/3 in the ratio 30/45/25. Mass spectrometry of the mixture also confirmed the formation of **3**: HRMS (ESI): Calcd for C₈H₁₅N₂O₂S₂ [M–H]⁺: *m/z* = 235.05695, found: *m/z* = 235.05702.

3. Results

3.1. Cyclisation of Homocysteine: Experimental Results

We tested the behavior of **1** in water at various pH values and temperatures. Results are presented in Figure 1. Ratios were estimated using ¹H NMR spectroscopy. An example of the obtained spectrums is presented in Figure 2.

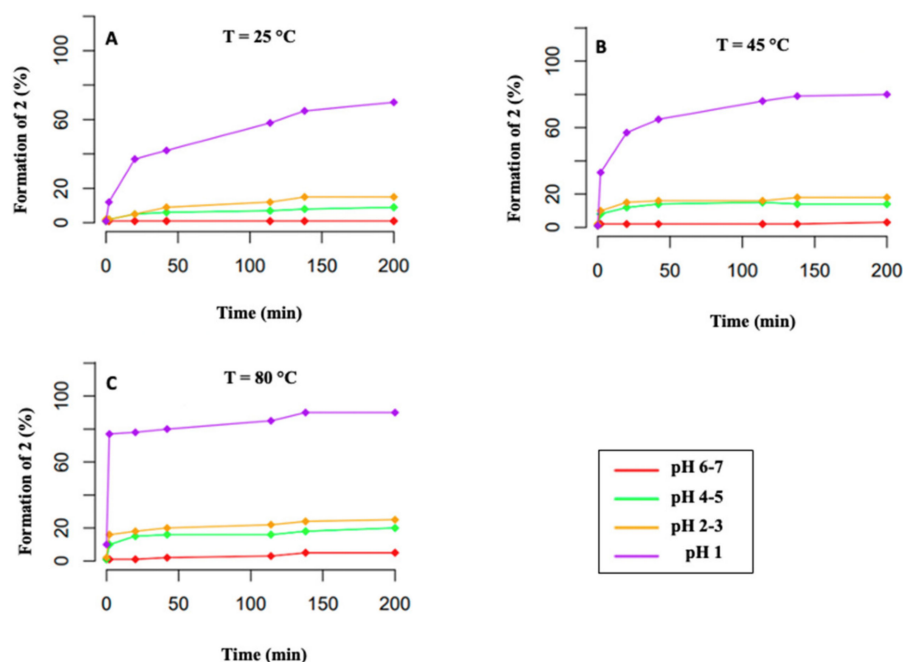


Figure 1. Formation of thiolactone **2** from **1** at various pH values and temperatures. (A), 25 °C; (B), 45 °C; (C), 80 °C.

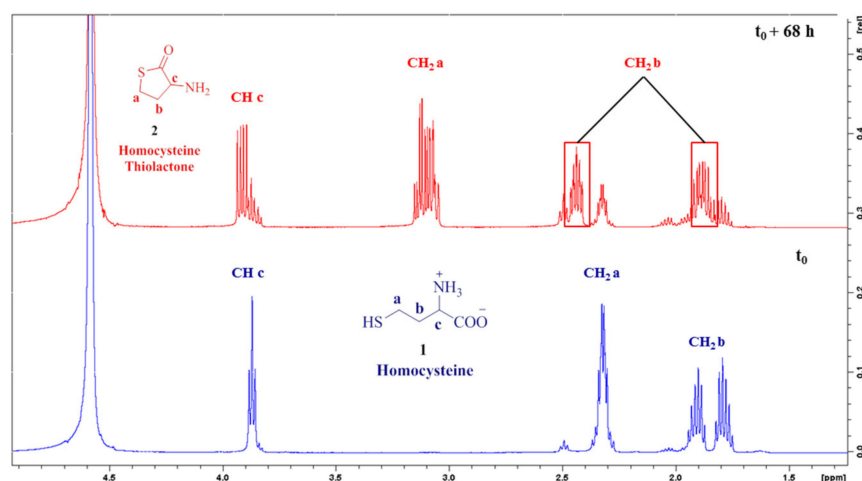
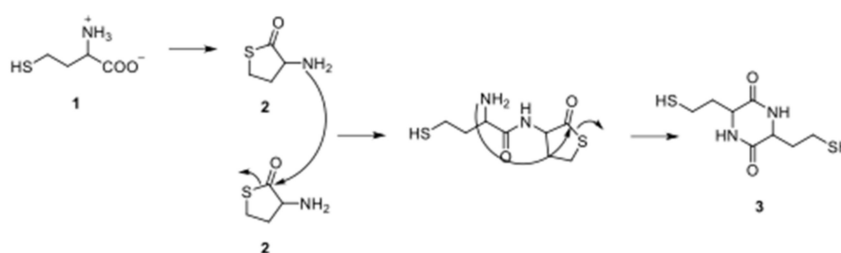


Figure 2. ^1H NMR of **1** (blue) and its transformation into **2** (after 68h, pH 1).

As observed in Figure 1c, at 80 °C, under very acidic conditions (pH 1), the dominant species is the thiolactone **2** (ratio **1/2**: 15/85). At this same temperature, under less acidic conditions (pH 2 to 5), when the equilibrium is reached, the amount of **2** is still important (ca. 15–20%). Only at nearly neutral pH does the proportion of **2** become smaller (a few percent). Results at lower temperatures (45 and 25 °C) are similar.

To mimic primitive conditions that may have existed, for instance, on warm beaches or near volcanoes, we also tested drying conditions [18]. A priori, such conditions should be favorable for water elimination reactions. Water was evaporated from an aqueous solution of **1** over silica, and the resulting dry solid was left at 80 °C overnight. The water-soluble content of this solid was recovered in D_2O . The ^1H NMR of the recovered mixture showed it to consist of a mixture of **1**, **2** and piperazinedione **3** [19] (Scheme 2). The **1/2/3** ratio was 30/45/25. Piperazinedione **3**, the cyclic form of HcyHcy, is a dimer of **1**, whose formation can be explained only via the opening of **2**. Thus, **2** can form not only in acidic water but also on a dry surface. As it was obtained here with a high concentration, two molecules of **2** reacted together, which means that the thiolactone underwent amide bond formation, **3** being a cyclic dipeptide.



Scheme 2. Mechanism of the dimerization of homocysteine **1**.

Additionally, when we dissolved preformed **2** in water at pH 7, its hydrolysis was slow even at 80 °C (less than 1% hydrolysis after 2 weeks). Only under basic conditions was the opening of **2** quicker (60% hydrolysis after 24 h at 25 °C).

3.2. Formation of Dipeptides from Homocysteine Thiolactone and Its *N*-formyl and *N*-acetyl Derivatives

We then tested the reactivity of **2** with amino acid. All reactions were run at 45 °C. Representative experiments are summarized in Figure 3 (other examples, including examples with small peptides, are presented in Supplementary Materials). At pH values lower than 7, no reaction was observed. At pH 7 (fixed by the addition of imidazole to the reaction mixture) and 8 (NaHCO_3) all tested reactions were efficient, with conversions to dipeptides being a little better at pH 8.

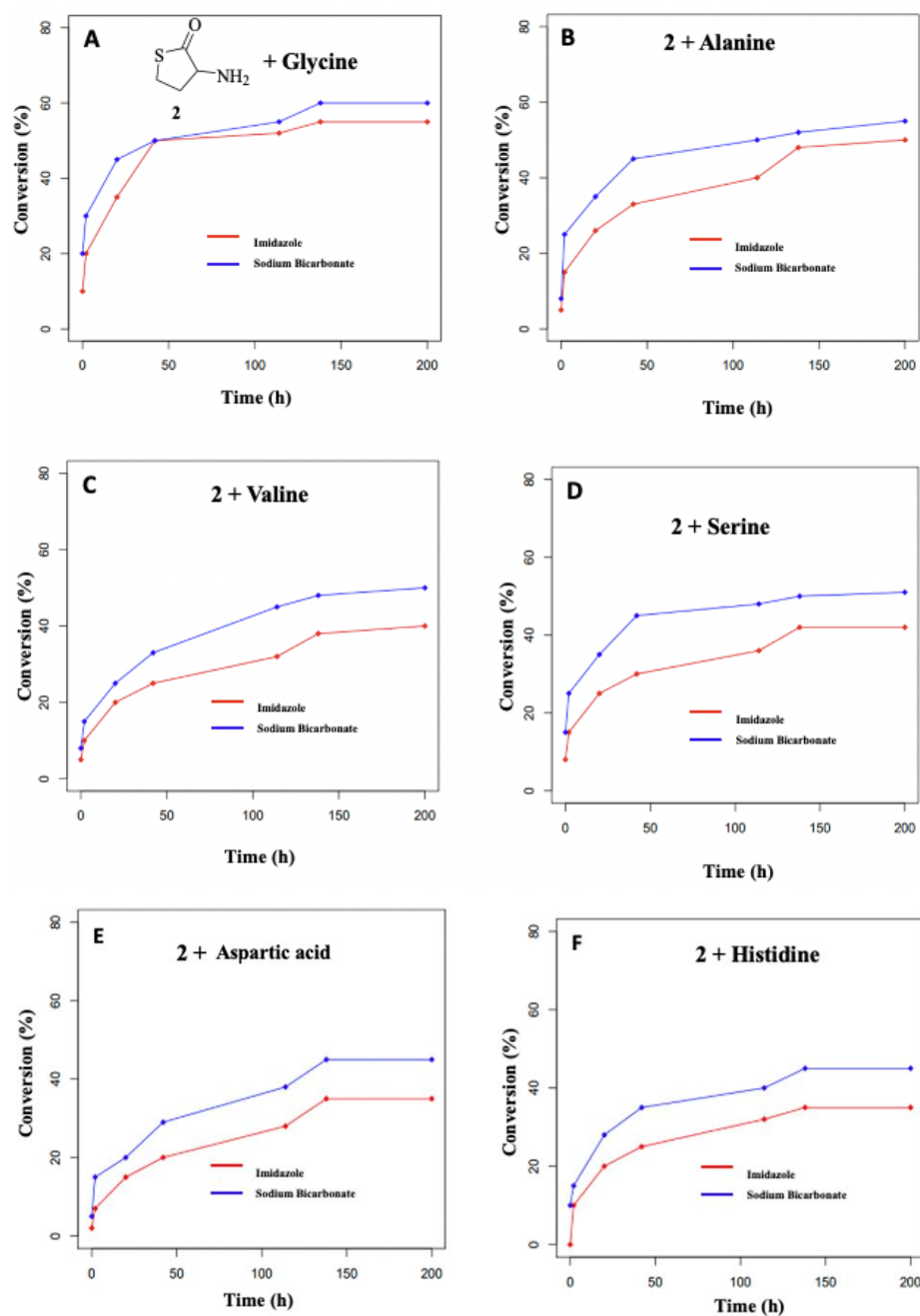


Figure 3. Evolution of the reaction of thiolactone **2** with various amino acids at pH 7 (red) and 8 (blue). % Conversion to dipeptide determined by ^1H NMR. (A), with glycine; (B), with alanine; (C), with valine; (D), with serine; (E), with aspartic acid; (F), with histidine.

We also tested the reactivity of the formylated thiolactone **4** with glycine (Figure 4). The reaction was run at 45 °C, pH 8, for 24 h, after which time the ^1H NMR of the reaction mixture clearly indicated the formation of **5**.

Similarly, *N*-acetyl homocysteine thiolactone **6** reacted with amino acids. Representative examples are presented in Figure 5 (other examples in Supplementary Materials).

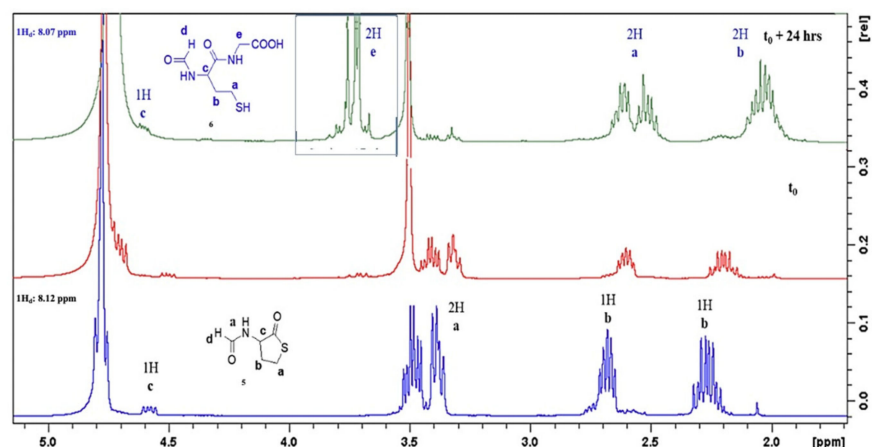


Figure 4. ^1H NMR monitoring of the reaction of **4** with glycine. Blue: **5**. Red: reaction mixture at t_0 . Green: reaction mixture after 24 h, showing the formation of the expected dipeptide **5**.

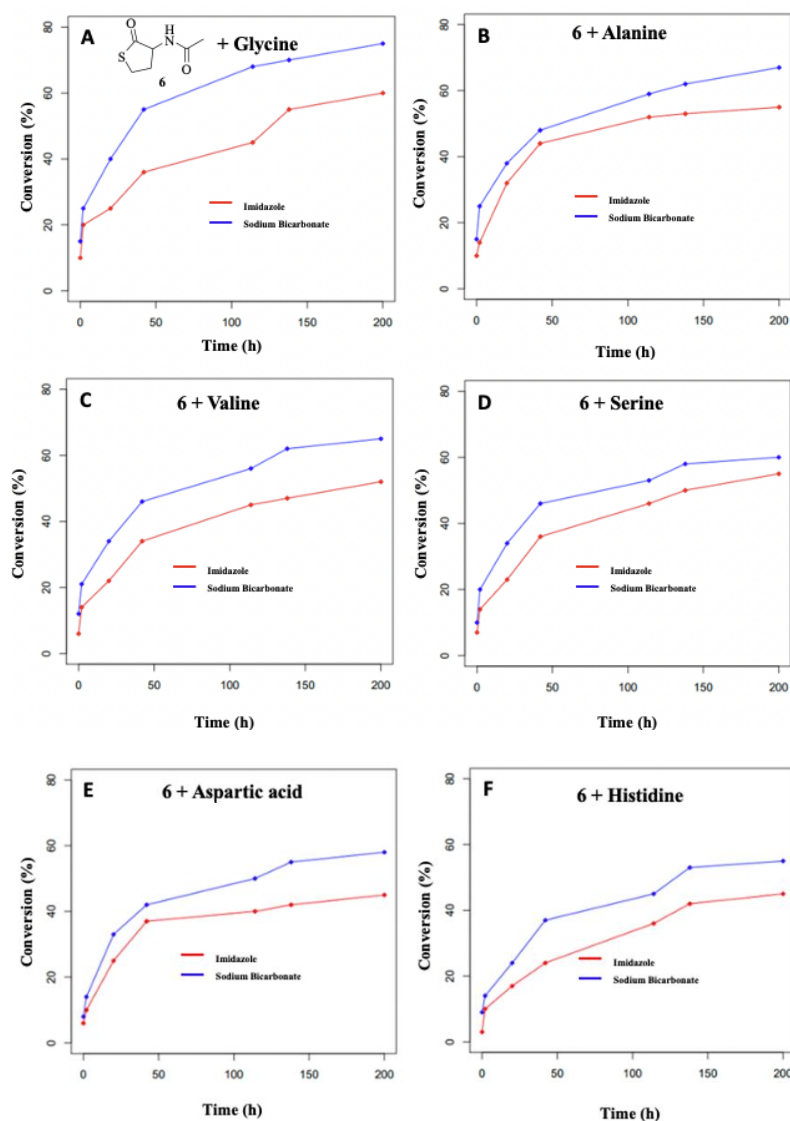


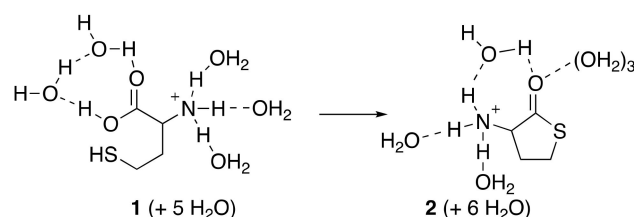
Figure 5. Evolution of the reaction of *N*-acetyl homocysteine thiolactone **6** with various amino acids at pH 7 (red) and 8 (blue). % Conversion to dipeptide determined by ^1H NMR. (A), with glycine; (B), with alanine; (C), with valine; (D), with serine; (E), with aspartic acid; (F), with histidine.

4. Discussion

4.1. Cyclisation of Homocysteine: Theoretical Calculations

Using DFT calculations, both static and dynamic, mechanical insights into the cyclisation can be proposed. All the static calculations were performed using Gaussian 16 [20] at the B3LYP level with the def2-TZVP basis set [21] and solvent effects using the SMD (Solvent Model Density) [22] option to mimic the bulk effect of water. Dispersion was added through the D3 version of Grimme with Becke–Johnson damping. The comparison of the results with and without D3BJ is given in the Supplementary Materials [23–25]. All the optimizations were followed by a frequency calculation to check the nature (minimum or transition state) of the optimized geometry. We also added explicit water molecules. For example, three water molecules are located around the NH_3^+ group.

The global energetical cost was computed using the reaction given on Scheme 3. The reaction is slightly endo-energetic by around $+3 \text{ kcal}\cdot\text{mol}^{-1}$ (See Supplementary Materials for more information and different levels). However, if we used $\text{H}_3\text{O}^+\text{-H}_2\text{O}$, around the carboxylic acid in Scheme 3 instead of two water molecules, the reaction is slightly exothermic with a ΔH of around $-7.35 \text{ kcal}\cdot\text{mol}^{-1}$, which is coherent with a reaction more efficient in acidic condition.



Scheme 3. Cyclisation of solvated homocysteine 1.

To clarify the discussion of the mechanism, it is important to specify some geometrical characteristics of the reactant and the product. For the reactant, the length of the SH bond was around 1.35 \AA , the final C-S bond length was around 1.80 \AA against 3.3 \AA in the starting nonlinear conformer. A picture of the dynamics of **1** in water and of the surrounding hydrogen bonds network can also be useful. Therefore, we used a General Amber Force Field (GAFF) [26], using Amber18 [27] to describe **1** and run a dynamic of **1** plus 2410 water molecules and one chlorine anion for neutrality for 70 ns after a process of equilibration (see Supplementary Materials for details). The trajectories were analyzed using cpptraj, the program-processing amber trajectories files [28]. We then extracted data on hydrogen bonds and distances. To define the existence of a hydrogen bond, the default values of cpptraj were always used: i.e., the distance cutoff from acceptor to donor heavy atom was 3.00 \AA and the angle cutoff was set to 135° . For the hydrogen bond surrounding the SH group, a cutoff of 3.4 \AA was used and the angle cutoff maintained at 135° . The OH group of the carboxylic acid function was then involved as a donor of hydrogen bond with water 96% of the time of the simulation with an average distance between the two oxygens involved of 2.687 \AA and an average angle of 163° . The three hydrogens of the NH_3^+ group were also involved in a hydrogen bond 74% of the time of the simulation with an average distance of 2.843 \AA and an average angle of 157° . The thiol function acted as a donor of hydrogen bonds toward water 58% of the time of the simulation with an average sulfur-oxygen bond distance of 3.199 \AA and an average angle of 154° . Then the oxygen of the carbonyl function of the carboxylic group was an acceptor of hydrogen bonds 52% of the time of the simulation (2.809 \AA and 156° for the average characteristics). The oxygen of the OH function of the carboxylic acid acted as an acceptor of hydrogen bonds 20% of the time of the simulation (2.841 \AA and 155°). The carbon–sulfur bond to be made during the cyclisation ranged from 3 to 6 \AA . The histogram from Figure 6 shows that the C-S bond distance is indeed very short, between 3 and 3.5 \AA , 5% of the time of the simulation, and 20% of the time the C-S bond distance ranged from 3 to 4 \AA . The bond distance reached its highest probability between 4.5 and 5 \AA , a range of distance observed nearly 40% of the

time during the simulation. These data are shown on Figure 6 and confirm that the folded conformation needed for the cyclisation of **1** is likely in water.

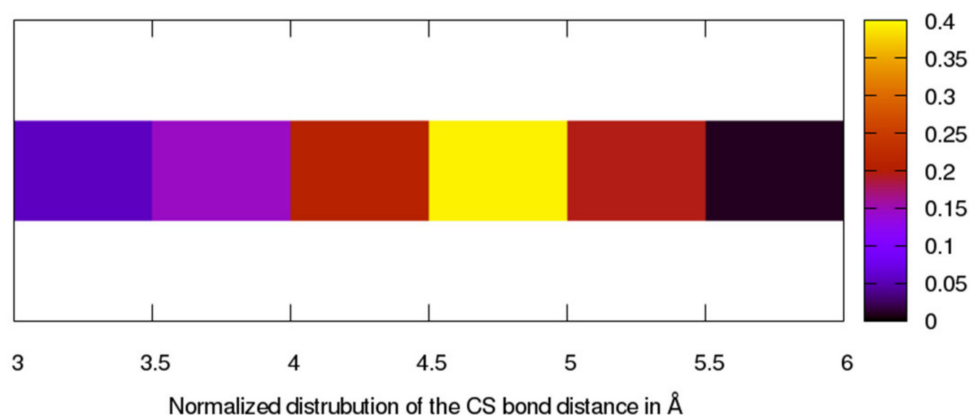
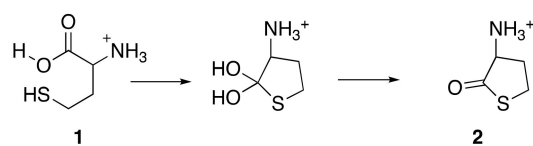


Figure 6. Normalized distribution of the carbon–sulfur bond distance. The total probability over all the bins is 1.

Due to the water network and the different proton transfers involved in the reaction, we decided to use QM/MM (quantum mechanics/molecular mechanics) calculations and metadynamics [29] using the CP2K software [30]. The setup of the system is described in the Supplementary Materials. The QM part consists of the organic molecule plus the surrounding water molecules at a distance of around 6 Å, i.e., 79 water molecules including a hydronium cation. Using this tailored quantum part, made up of the organic molecule and a reasonable number of water molecules around it, gives the protons the opportunity to move from homocysteine **1** to the solvent while remaining close to **1**. The objective is to keep the “acidic” aspect of the solution close to the molecule. Metadynamics was used to overcome the problem of observing rare events in conventional molecular dynamics and of finding the reaction coordinate. A series of small repulsive Gaussian potentials centered on the values of some collective variables (CV) were added during the dynamics, preventing the system from revisiting the same points in configurational space and creating a history-dependent multidimensional-biasing potential. A time step of 0.5 fs was used for the dynamics, and the hills of 1 kcal·mol^{−1} were added every 20 fs.

To study this reactivity, three collective variables were used. The first one was the carbon sulfur bond to be formed; the second was the S-H to be broken; the C-O bond was the last variable. The width of the Gaussian used was 0.3 bohr, and a reflective wall at 6.2 bohr was used to prevent sampling the C-S bond at large values, which is of no interest since we study the cyclisation. The observed reaction was the creation of the C-S bond concomitant to the proton abstraction from the S-H bond. The C-O bond was not broken. However, its elongation led to its protonation, and a diol intermediate was then identified (Scheme 4). The barrier needed is around 25 kcal·mol^{−1}. The joint formation of the CS bond and the breaking of the SH bond during the metadynamics can be seen on Figure 7. The limitation of the CS bond to 3.5 Å is due to the addition of a wall to prevent the useless sampling of these conformations during the metadynamics.



Scheme 4. A transient diol is formed during the cyclization process.

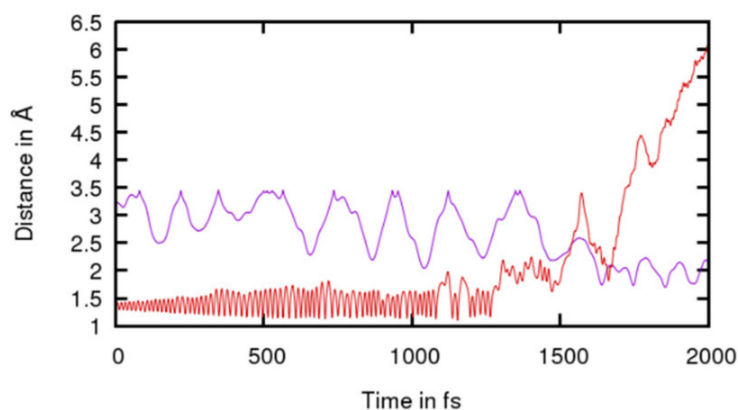


Figure 7. Evolution during the metadynamics of the S-H bond distance in red and of the C-S bond distance in purple in Å as a function of the time in fs.

Returning to the static calculation at the B3LYP/def2TZP level, the diol (Scheme 4) plus two water molecules and with the three water molecules around NH_3^+ lies $12.5 \text{ kcal}\cdot\text{mol}^{-1}$ higher in energy (ΔG of $16.0 \text{ kcal}\cdot\text{mol}^{-1}$) than the starting molecule. A proton transfer can occur using the two water molecules as a proton relay leading to the final cyclic molecule. The barrier needed is of $15.1 \text{ kcal}\cdot\text{mol}^{-1}$ (ΔG^\ddagger of $14.1 \text{ kcal}\cdot\text{mol}^{-1}$).

Thus, we propose a two-steps mechanism, with the first leading to a diol $12.5 \text{ kcal}\cdot\text{mol}^{-1}$ higher in energy than homocysteine with a barrier of $25 \text{ kcal}\cdot\text{mol}^{-1}$. Then, through a barrier of $15 \text{ kcal}\cdot\text{mol}^{-1}$, the diol leads to the thiolactone.

4.2. Hypothesis: Hcy Was the First StartResidue

We have recently reported that in a “nitrile world” (cyanosulfidic world [4]) **2** was a probable intermediate in the synthesis of homocysteine **1** itself [11]. From what is reported above (both experimentally and theoretically), it is now clear that even not using nitriles, if **1** was present, then **2** was also. Highly acidic conditions are particularly favorable for the formation of **2**. Such conditions could have been present on the early Earth, for instance near black smokers and volcanoes. Under such conditions the cyclized form **2** indeed dominated the linear form **1**. However, even under less acidic conditions, **2** was present in a reasonable proportion. Due to the huge amount of CO_2 present in the atmosphere, the primitive global ocean was probably slightly acidic [31], meaning that in this ocean, the ratio between **1** and **2** was ca. 5 to 1. Even under nearly neutral conditions, at least if the temperature was high enough (which was probably the case [31]), some quantities of **2** were present. Furthermore, we have noticed that, once formed, **2** is stable in neutral water, even at high temperatures. This is consistent with the known stability of **2** in human serum [32]. Of course, under basic conditions, its hydrolysis is faster. However, its half-life at pH ca. 8 remains non-negligible. We have shown that such pH conditions (from 7 to 8) are necessary to obtain good yields of dipeptides from the reaction of **2** with amino acids.

From these results it appears that the prebiotic synthesis of homocysteine dipeptides would require a pH gradient from acidic (maybe even strongly acidic) for the formation of **2** to neutral or slightly basic for its opening by an amino acid. The pH gradients are of major importance in biology, and they are quite usual in nature, for instance near volcanoes or black smokers [33–35]. There is therefore no particular difficulty in postulating their use in prebiotic chemistry [36–38].

It is thus reasonable to postulate the existence of Hcy-AA dipeptides on the early Earth (where AA stands for any amino acid), and as homocysteine was able to self-activate (forming its thiolactone), these dipeptides could well have been overrepresented compared to the other possible $\text{AA}_1\text{-AA}_2$ pairs (for which a necessarily slower inter-molecular activation of AA_1 was required) [39–41]. We hypothesize that these overrepresented Hcy-AA interacted with RNAs (proto-tRNAs) to form Hcy-AA-RNAs, and that the presence of a Met residue at the start of the synthesis of all actual proteins is a consequence of these

primitive reactions. If so, the very first “start residue” was not Met but Hcy, and it was not coded: there was no starting codon. It is only later, when homocysteine was excluded from the list of the 20 canonical amino acids (probably because its presence was detrimental to the stability and to the properties of proteins [42]) that it was replaced as first residue by its S-methylated analogue, methionine, which implied the introduction of AUG as start codon. The fact that AUG is not always the codon for the start residue Met can then be regarded as a consequence (and a remembrance) of this late substitution of Hcy by Met, with the need to introduce a start codon corresponding to Met where there was none before. In any case, this would be consistent with the fact that an amino acid that was introduced so lately in the genetic code could still be chosen as the start residue. Moreover, the presence of Hcy residues elsewhere in peptides would have allowed the development of the typical reactivity of thiols in addition to that of Cys residues. Even thiolactone **2** could have been involved in varied radical processes [43,44].

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/biom12040555/s1>, Representative NMR spectra, Supplementary examples of dipeptide formation, Details about theoretical calculations. References [23,26,27,30,45–48] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, Y.V.; methodology, Y.V.; validation, Y.V., A.M. and S.Y.-S.; experimental investigation, S.Y.-S.; theoretical calculations, A.M.; writing—original draft preparation, Y.V., A.M., S.Y.-S.; writing—review and editing, Y.V.; supervision, Y.V.; project administration, Y.V.; funding acquisition, Y.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Agence Nationale de la Recherche in the framework of the Investissements d’Avenir Program (ANR-15-IDEX-02), through the funding of the “Origin of Life” project of the Université Grenoble Alpes, Labex ARCANÉ, and CBH-EUR-GS (ANR-17-EURE-0003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We acknowledge ICMG—Plateforme Chimie Nanobio for mass spectrometry.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Krishnamurthy, R. Giving Rise to Life: Transition from Prebiotic Chemistry to Protobiology. *Acc. Chem. Res.* **2017**, *50*, 455–459. [[CrossRef](#)] [[PubMed](#)]
2. Zahnle, K.; Schaefer, L.; Fegley, B. Earth’s Earliest Atmospheres. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a004895. [[CrossRef](#)] [[PubMed](#)]
3. Muchowska, K.B.; Varma, S.J.; Moran, J. Nonenzymatic Metabolic Reactions and Life’s Origins. *Chem. Rev.* **2020**, *120*, 7708–7744. [[CrossRef](#)] [[PubMed](#)]
4. Sutherland, J.D. The Origin of Life—Out of the Blue. *Angew. Chem. Int. Ed.* **2016**, *55*, 104–121. [[CrossRef](#)]
5. Marchi, S.; Bottke, W.F.; Elkins-Tanton, L.T.; Bierhaus, M.; Wuennemann, K.; Morbidelli, A.; Kring, D.A. Widespread mixing and burial of Earth’s Hadean Crust by asteroid impacts. *Nature* **2014**, *511*, 578–582. [[CrossRef](#)]
6. Youssef-Saliba, S.; Vallée, Y. Sulfur Amino Acids: From Prebiotic Chemistry to Biology and Vice Versa. *Synthesis* **2021**, *53*, 2798–2808. [[CrossRef](#)]
7. Jakubowski, H. Homocysteine Editing, Thioester Chemistry, Coenzyme A, and the Origin of Coded Peptide Synthesis. *Life* **2017**, *7*, 6. [[CrossRef](#)]
8. Perla-Kajan, J.; Twardowski, T.; Jakubowski, H. Mechanisms of Homocysteine Toxicity in Humans. *Amino Acids* **2007**, *32*, 561–572. [[CrossRef](#)]
9. Parker, E.T.; Cleaves, H.J.; Callahan, M.P.; Dworkin, J.P.; Glavin, D.P.; Lazcano, A.; Bada, J.L. Prebiotic Synthesis of Methionine and Other Sulfur-Containing Organic Compounds on the Primitive Earth: A Contemporary Reassessment Based on an Unpublished 1958 Stanley Miller Experiment. *Orig. Life Evol. Biosph.* **2011**, *41*, 201–212. [[CrossRef](#)]
10. Shalayel, I.; Vallée, Y. Chemistry of Homocysteine Thiolactone in A Prebiotic Perspective. *Life* **2019**, *9*, 40. [[CrossRef](#)]
11. Deepak, R.N.V.K.; Chandrakar, B.; Sankararamkrishnan, R. Comparison of Metal Binding Strength between Methionine and Cysteine Residues: Implications for the Design of Metal-Binding Motifs in Proteins. *Biophysic. Chem.* **2017**, *224*, 32–39. [[CrossRef](#)] [[PubMed](#)]

12. Kozłowski, L.P. Proteome-pl: Proteome Isoelectric Point Database. *Nucleic Acids Res.* **2017**, *45*, 1112–1116. [[CrossRef](#)] [[PubMed](#)]
13. Trifonov, E.N. The Triplet Code from First Principles. *J. Biomol. Struct. Dyn.* **2004**, *22*, 1–11. [[CrossRef](#)] [[PubMed](#)]
14. Blattner, F.R.; Plunkett, G., III; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; et al. The Complete Genome Sequence of Escherichia coli K-12. *Science* **1997**, *277*, 1453. [[CrossRef](#)] [[PubMed](#)]
15. Belinky, F.; Rogozin, I.B.; Koonin, E.V. Selection on Start Codons in Prokaryotes and Potential Compensatory Nucleotide Substitutions. *Sci. Rep.* **2017**, *7*, 1242. [[CrossRef](#)]
16. De Duve, C. A Research Proposal on the Origin of Life. *Orig. Life Evol. Biosph.* **2003**, *33*, 559–574. [[CrossRef](#)] [[PubMed](#)]
17. Risse, C. Homocysteinthiolacton-Abkömmlinge und Verfahren zu deren Herstellung. German Patent DE1972-2213028, 21 November 1985.
18. Abelson, P.H. Chemical Events on the Primitive Earth. *Proc. Natl. Acad. Sci. USA* **1966**, *55*, 1365–1372. [[CrossRef](#)]
19. Mendham, A.P.; Spencer, J.; Chowdhry, B.Z.; Dines, T.J.; Mujahid, M.; Palmer, R.A.; Tizzard, G.J.; Coles, S.J. X-Ray Crystallographic Structure and Absolute Configuration of the Cyclic Di-amino Acid Peptide: Cyclo(L-HomoCySH-L-HomoCySH). *J. Chem. Crystallogr.* **2011**, *41*, 1328–1334. [[CrossRef](#)]
20. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16, Revision B.01*; Gaussian, Inc.: Wallingford, CT, USA, 2016.
21. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Phys.* **2006**, *8*, 1057–1065. [[CrossRef](#)]
22. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tension. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [[CrossRef](#)] [[PubMed](#)]
23. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H.A. Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)] [[PubMed](#)]
24. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [[CrossRef](#)] [[PubMed](#)]
25. Johnson, E.R.; Becke, A.D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124*, 174104. [[CrossRef](#)] [[PubMed](#)]
26. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)]
27. Case, D.A.; Ben-Shalom, I.Y.; Brozell, S.R.; Cerutti, D.S. *AMBER 2018*; University of California: San Francisco, CA, USA, 2018.
28. Roe, D.R.; Cheatham, T.E. *PTRAJ and CPPTRAJ*: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [[CrossRef](#)]
29. Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566. [[CrossRef](#)]
30. Kühne, T.D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V.V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R.Z.; Schütt, O.; Schiffrmann, F.; et al. CP2K: An Electronic Structure and Molecular Dynamics Software Package—Quickstep: Efficient and Accurate Electronic Structure Calculations. *J. Chem. Phys.* **2020**, *152*, 194103. [[CrossRef](#)]
31. Pinti, D.L. The Origin and Evolution of the Oceans. *Lect. Astrobiol.* **2005**, *1*, 83–112. [[CrossRef](#)]
32. Chwatko, G.; Jakubowski, H. The Determination of Homocysteine-thiolactone in human plasma. *Anal. Biochem.* **2005**, *337*, 271–277. [[CrossRef](#)]
33. Velázquez-Ríos, I.O.; Rincón-Rosales, R.; Gutiérrez-Miceli, F.A.; Alcántara-Hernández, R.J.; Ruíz-Valdiviezo, V.E. Prokariotic diversity across a pH gradient in the « El Chichón » crater-lake: A naturally thermo-acidic environment. *Extremophiles* **2022**, *26*, 8. [[CrossRef](#)] [[PubMed](#)]
34. Butler, J.B.; Nesbitt, R.W. Trace element distributions in the chalcopyrite wall of a black smoker chimney: Insights from laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS). *Earth Planet. Sci. Lett.* **1999**, *167*, 335–345. [[CrossRef](#)]
35. Keil, L.M.R.; Möller, F.M.; Kieß, M.; Kudella, P.W.; Mast, C.B. Proton gradient and pH oscillations emerge from heat flow at the microscale. *Nat. Commun.* **2017**, *8*, 1897. [[CrossRef](#)] [[PubMed](#)]
36. Hudson, R.; de Graaf, R.; Strandoo Rodin, M.; Ohno, A.; Lane, N.; McGlynn, S.E.; Yamada, Y.M.A.; Nakamura, R.; Barge, L.M.; Braun, D.; et al. CO₂ reduction driven by a pH gradient. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 22873–22879. [[CrossRef](#)]
37. Lane, N. Proton gradients at the origin of life. *Bioessays* **2017**, *39*, 1600217. [[CrossRef](#)]
38. Holm, N.G.; Hennes, R.J.C. Hydrothermal systems: Their varieties, dynamics, and suitability for prebiotic chemistry. *Orig. Life Evol. Biosph.* **1992**, *22*, 15–31. [[CrossRef](#)]
39. Foden, S.; Islam, C.; Fernández-García, C.; Maugeri, L.; Sheppard, T.D.; Powner, M.W. Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water. *Science* **2020**, *370*, 865–869. [[CrossRef](#)]
40. Canavelli, P.; Islam, S.; Powner, M.W. Peptide ligation by chemoselective aminonitrile coupling in water. *Nature* **2019**, *571*, 546–549. [[CrossRef](#)]
41. Sauer, F.; Haas, M.; Sydow, C.; Siegle, A.F.; Lauer, C.A.; Trapp, O. From amino acid mixtures to peptides in liquid sulphur dioxide on early Earth. *Nat. Commun.* **2021**, *12*, 7182. [[CrossRef](#)]
42. Jakubowski, H. Homocysteine modification in protein structure/function and in human disease. *Physiol. Rev.* **2019**, *99*, 555–604. [[CrossRef](#)]

43. Fucks, E.; Huber, L.; Schinkel, T.; Trapp, O. Investigation of straightforward, photoinduced alkylations of electron-rich hetero-compounds with electron-deficient alkyl bromides in the sole presence of 2,6-lutidine. *Eur. J. Org. Chem.* **2020**, *2020*, 6192–6198. [[CrossRef](#)]
44. Closs, C.; Bechtel, M.; Trapp, O. Dynamic exchange of substituents in a prebiotic organocatalyst: Initial steps towards an evolutionary system. *Angew. Chem. Int. Ed.* **2022**, *61*, e202112563. [[CrossRef](#)] [[PubMed](#)]
45. Bayly, C.I.; Cieplak, P.; Cornell, W.; Kollman, P.A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280. [[CrossRef](#)]
46. Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. Cp2k: Atomistic Simulations of Condensed Matter Systems. *WIREs Comput. Mol. Sci.* **2014**, *4*, 15–25. [[CrossRef](#)]
47. Van de Vondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* **2007**, *127*, 114105. [[CrossRef](#)]
48. Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710. [[CrossRef](#)]