

# Derivation and Progress of Non-enzymatic RNA Replication

Xinyu Wu

Beijing Jingshan School Caofeidian Branch International Department, Tangshan, China

Email: kellywuxinyu@gmail.com

**Abstract**—Molecular replication has branched into two parts, on one hand, it involves the non-informational copying of polymers which provides a replication environment indirectly, and on the other hand, the informational copying, including DNA or RNA replication, is a direct replication system. On account of the central dogma of molecular biology, after the transcription from DNA to mRNA, some living organisms require a specific RNA polymerase to catalyze and control the reaction of RNA self-replication. Based on the RNA world hypothesis, the self-replication of RNA at the earlier stage of the origin of life occurred under an enzyme-free condition. RNA replication process needs three steps: (a) separation of the duplex, (b) monomer extension of each strand, and (c) the end to end propagation of RNA strand. It was suggested that non-enzymatic RNA copying could take place in the presence of metal ions, like  $Mg^{2+}$  or  $Fe^{2+}$ . Strand separation can be achieved by pH or temperature oscillation, and the RNA monomer extension could be attainable with the involvement of two activated nucleotides. This review gives a brief summary of the accomplishments scientists have made, and also concerns about several impediments in this process as well as their solutions.

**Index Terms**—divalent ions, imidazolium dinucleotide, RNA world, the origin of life

## I. INTRODUCTION

In light of the origin of life, the building blocks of earlier life were developed in an oxygen-lacking environment, and in the reducing atmosphere with nitrogen ammonia and methane [1]. A common belief states that life emerged accidentally from a mixture of compounds by a simple autocatalyst process. The cyanosulfidic chemistry and hydrothermal vents are supposed to be the examples that generate new polymers and minerals from the aggregation of already existed molecules. Under the condition of the primitive earth, the organic compounds that could be used in the abiotic synthesis are not abundant; a vast majority of them are organic compounds with a small amount of carbon like oligonucleotides, whose formation involves electric discharge, UV radiation, and shock waves [2].

Amino acids are supposed to be the easiest biomolecule to synthesis. It reacts with the molecules in the system and form peptides. As in amide formation the leaving group is hard to attack back due to the strength of

C-N bonding, the condensation of two amino acids is faster than its hydrolysis. This property ensures the successful synthesis of the amino acids, which provides the non-enzymatic replication with its intermedia--imidazolium dinucleotide.

RNA is supposed to be an extraordinarily sophisticated compound. This indicates the RNA synthesis is a reaction that requires a relatively massive amount of enzymes' participation. The RNA predecessor in the condition of the origin of life was different from what the RNA is now. Furthermore, it might not be affected by the  $OH^-$  group in its wide range of reactions [3]. However, based on the RNA world hypothesis, the RNA strand plays both roles in either the transcription and its self-replication in the origin of life. It acts as both a template and a catalyst without the specific polymerase to initiate the Darwin evolution. A vast amount of enzymes are made of protein, but we do have some that are composed of RNA. The most typical example is the ribozyme, a kind of RNA molecules that may accelerate the chemical reaction. In most reactions aimed at abandoning the catalysts, the presence of divalent metal ions is necessary. But it has been suggested that alternatively, we can allow the reactions to occur by adding high concentrated monovalent ions, or cobalt(III) hexamine complex( $CoH_{18}N_6$ ). It has been stated that this discovery can lead to the conclusion that cations are functioning at the neutralisation of anions in phosphodiester backbones and increase its stability. The charged ions could alter the pKa and nucleophilicity of its adjacent functional group and consequently being involved in this reaction [4].

The separation of RNA strand under the condition of primitive earth allows the informational copying of amino acids, the strand unzipped and followed by the extension from the primer. A specific RNA polymerase is required in this step, however, as in the origin of life, protocells are developed without the assist of enzymes, so consequently, recent researchers are eager to find out the procedure to reemerge the non-enzymatic RNA replication.

Szostak's group demonstrated that the enzyme-free primer extension of an RNA duplex can be established under a dinucleotide mediated addition reaction. The reaction can be observed by X-rays under frozen condition. They introduced the RNA nucleotides with crystallised RNA backbone complex. Eventually, the nucleotides on the RNA template are capable of joining

their complementary nucleotides in the solution. The nucleotides react with each other and form the intermediate. This intermediate then attaches to the primer ultimately and allows the primer to bind with one of its nucleotides, and then the other nucleotide will generate for further reactions [5].

Even though various ways have been revealed throughout contemporary research to allow the RNA replicating in an enzyme-free condition, the process is not entirely gorgeous. The unique structure of the RNA strand exposes several issues in the hypothesis of the RNA world. We have to take account of the rate to let the reaction occur. The RNA within organisms exists as a single strand, under basic conditions, the chains will break so that it is less stable than DNA. However, due to the effect of depurination, under the acidic conditions, the purine bases in DNA double helix will drop of the pentose sugar, and the DNA structure is broken simultaneously. Under this circumstance, RNA is supposed to be more stable than DNA. However, in the primitive earth environment, the RNA exists as double helix joined together by the intermolecular force made by hydrogen bonding. The hydrogen bonds attach the fragments tightly in a large extent of polar molecules like alcohol, as well as nucleotides. We can find out with evidence that RNA duplex is significantly stable than DNA so that inversely it requires higher enthalpy to break the RNA helix. According to the RNA world model and the structure of RNA chains, after the replication in the formation of RNA duplex, as the accumulation effect of hydrogen bonding, it forms a dead-ended helix and loses its function as a catalyst consequently.

## II. DISCUSSION

### A. Metal Ions Catalysed Non-enzymatic RNA Strand Formation and Its Folding

The metal ions are extensively used in contemporary biology. It maintains the concentration gradient in metabolism and plays a role in the structure of the protein. The wide range of metal ions and their various properties function in diverse areas.  $Mg^{2+}$ , acting as a counter to stabilise the enzymes in reactions releasing ATP like aerobic respiration and muscle contraction. Magnesium is the 8th most abundant element in the contemporary earth in the form of crust deposits. 99% of  $Mg^{2+}$  in the human body are found in bones muscles and soft tissues.  $Mg^{2+}$  deficiency may result in disorders in nerve systems and circulatory systems [6].  $Mg^{2+}$  is the lightest metal so that it is applied in the industry as the alternative to some dense, heavy metals as  $Mg^{2+}$  based alloy [7]. Due to the application of metal in modern society. Scientists are attempting to introduce metal ions in the origin of life.

$Fe^{2+}$  is commonly known for its property to carry oxygen and electrons and its good adaptation in both oxidised and reduced states, iron has been recognised as an essential factor in a wide range of chemical synthesis and medical treatments.  $Fe^{2+}$  might act as an electron donor in the oxygen-lacking environment, which made the presence of iron in the origin of life seems plausible.

Its concentration in the ocean could be regulated by the interaction between dissolved inorganic compounds with  $SiO_2$  [8]. And some research in geology reckons that the existence of  $Fe^{2+}$  in the origin of life is realisable due to the discovery of the deposit of  $Fe^{2+}$  in the Hamersley basin, this deposit is approximately 6 million years, and the ratio of  $Fe^{2+}$  to  $Mg^{2+}$  in this basin fits the calibration in an average rock which is 1.7 [9].

According to a paper published in 2019 by Szostak's group, the catalysis of  $Mg^{2+}$  in the spontaneous primer extension of RNA could avoid the hydrolysis of the already formed imidazolium-bridge. The imidazolium dinucleotide mediated model was formed by nucleophilic addition (Fig. 1), the hydroxyl group on the primer pentose sugar attack the next downstream nucleotide's 5' phosphate, followed by the electron delocalisation and makes the imidazolium becomes a proper leaving group. Such a reaction allows the imidazolium on the original nucleotide primer attach to the second downstream nucleotide and forms the intermedia that could promote the rate of non-enzymatic RNA primer extension as this bridge is chemically reactive. It forms a repetitive process as there will always be imidazolium groups formed as leaving group so it could catalyse the further reactions between the downstream nucleotides (Fig. 2) [10].

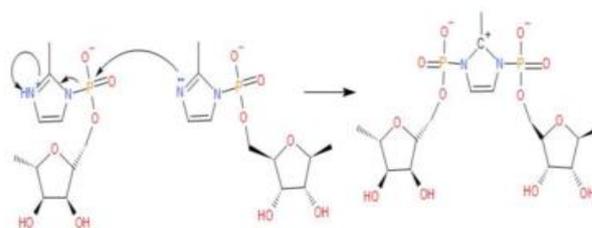


Figure 1. The mechanism of the formation of imidazolium dinucleotide intermedia.

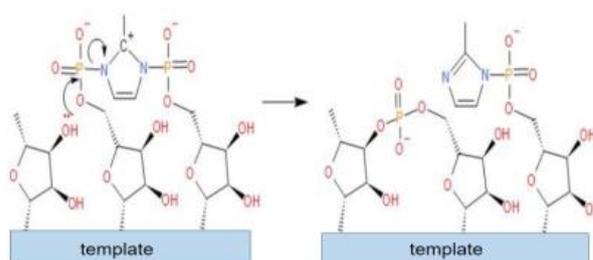


Figure 2. The primer extension of nucleotides along the RNA template with the assistance by the intermedia.

Since the templates of RNA were not capable of supporting the oligonucleotides with its desired concentration, scientists demonstrated an autocatalysed model using 4-thioUTP(4-thiouridine-5' triphosphate) (Fig. 3) and radiolabeled ATP, through the isolating of RNA ligase ribosome from a pool with 10 different RNA sequence. They controlled the RNA template in a shorter length to increase the extent of hybridisation. The RNA primer segments in the pool have been satisfied extended after eight rounds of selection as well as amplification. Moreover, results highly adaptable to long templates emerged by the round-18 ribosome, which is also efficient and precise in the turn-over the process of RNA

helix. By contrast, this method does not require the binding of the primer with the ribosome after the round tenth ribosome. Moreover, it successfully achieved the primer extension with multiple turns over rate. Also, the rate of reaction of primer extension for the Watson-Crick base pairs is supposed to be more efficient, with evidence it has higher fidelity in the extension of C-G rather than A-U [11].

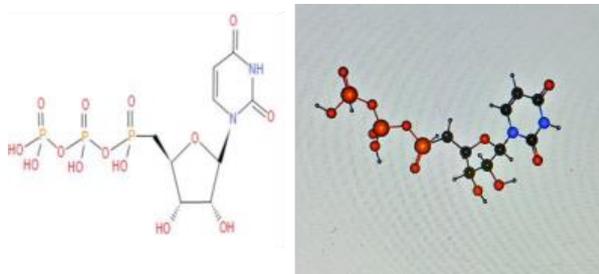


Figure 3. The structural formula of 4-thioUTP and its ball stick structure.

RNA folding is a process that happened simultaneously with RNA strand formation. The new structure that the RNA strand folded into could let it obtain specific functions and might bind with some other molecules. Together with the acidic and oxygen-free environment, divalent metal ions were present in the primitive earth, mainly in the pond and lakes. They are capable of taking over the role of certain catalysts like RNA ligase. Earlier research has shown that  $Mg^{2+}$  is available in the catalysing of the RNA strand folding, whereas Mg ions are redox inactive and have high charge density due to its small atomic radius. This brings  $Mg^{2+}$  with a high possibility to react with phosphate [12]. Such way using  $Mg^{2+}$  is not economically satisfied: while the  $Mg^{2+}$  catalyse the formation of RNA strand, it does catalyse its degradation which discomposes the polynucleotides by breaking their phosphodiester bond through hydrolysis, so that those facts make the whole reaction relatively slower than expected. Fortunately, via research, the substitutes of Mg ions have been found. The  $Fe^{2+}$  has two units less pKa than  $Mg^{2+}$  to promote the reaction occurs and is economically attainable. The  $Fe^{2+}$  can take over the role of  $Mg^{2+}$ , but for  $Fe^{3+}$ , it does not follow the same pattern, as it will lead to the precipitate of RNA molecules due to its strong affinity to react with phosphate, which is similar as  $Mg^{2+}$  as what we mentioned before.

#### B. Non-enzymatic Strand Separation of RNA by pH Changes and Temperature Oscillation

Based on the central dogma of molecular biology, when RNA translates into protein, the strand separation, or fork formation of RNA ensures the aggregation of nucleotides along with the RNA template. However, one RNA monostrand is able to act as a catalyst. The strong effect of hydrogen bonding forms a dead-ended product in the RNA duplex so that it entirely lost its function. It has been stated via the introduce of unpaired bases the melting point could be effectively lowered.

The involvement of even a slight amount of 2'-5' nucleotides backbones could reduce the energy intake to separate the RNA duplexes to a large extent, which is far more efficient than the 3'-5' backbones. As the participation of two unidentical backbones, i. e., both the 3'-5' and 2'-5' backbones are involved. That distorts the 3-dimensional structure of the template and made it seems imperfect and unstable; therefore, it requires less energy for separation of the RNA duplex [13].

A recent study in this area by Sutherland's group shown that to against the high melting point of RNA and enables it to undergo strand reannealing, with a circumstance related to the pH fluctuation (Fig. 4). As the denaturing of its active site, the temperature corresponding to different pH values could be modified by UV (ultraviolet radiation) and CD (circular dichroism spectrum) thermal studies as well as gel electrophoresis. Their study began by investigating the influence of the pH fluctuates when halving of the 13mer RNA molecule being denatured. Based on the data gained by the UV thermal melting curves, it was surprised that the melting point could be altered significantly under a pH value lower than 4.6. Besides, due to the process of protonation, the substances have increased acidity than previous and become more stable. Consequently, the CD curves were lower for the more acidic substances. So that a more acidic ambience was more likely to provide a plausible condition for RNA strand separation [14].

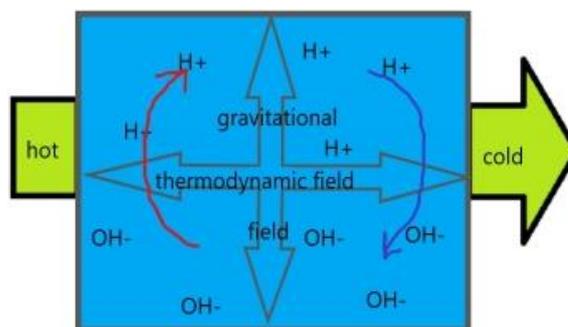


Figure 4. pH oscillation maintained by temperature change.

The pH value functioning in various areas in the origin of life, including the synthesis of nucleic acids and protein as well as their degradation, ion-sulfur peptide was a molecule which has high possibility to present in the origin of life. And it passes down electrons received from NADH (reduced form of nicotinamide-adenine dinucleotide) to hydrogen peroxide; the generation of hydroxyl ions by the reduction of the amount of hydrogen peroxide in the later phase of the origin of life helps the maintaining of the proton gradient in the primitive cell membrane [15]. Furthermore, keeping the proton gradient is essential in the degradation of RNA strand, as the proton gradient will ultimately be altered into an equilibrium state by diffusion, so pH oscillation has been suggested using heat flow to maintain this gradient.

A paper published in 2017 used a compartment called microfluidic device to resemble the condition in the

primitive earth, like the hydrothermal vents which was discovered in the Mid-Atlantic Ridge [16]. Their experiment obtained pH flux in the buffer solution by approximately 2 units by generating high convection velocity to offset and inhibit the molecules' diffusion to reach equilibrium. As the different temperature on two sides of the chamber, a horizontal thermodynamic field and a vertical gravitational field formed and causing the solution to rotate inside. A pH gradient is observed during this process, the gradient is maximum with the buffer solution of the smallest size and the most rapid diffusion rate. This method overcomes the lack of concentration in the research towards the origin of life and RNA strand separation [17].

Problems still exist at this stage. The strand separation of RNA is followed by the strand annealing, which requires a relatively neutral environment. [18] Nevertheless, the acidic surrounding at the phase of RNA separation cannot change instantaneously, which leads to an issue. Appropriate solutions towards this were still under study.

### C. Factors Impede the Non-enzymatic RNA Replication

The properties of the RNA strand and the condition in the primitive earth combined to make the non-enzymatic RNA replication sophisticated. To imitate the evolution of the protocell as well as maintain the efficiency of the reaction, scientists have attempted ways to tackle the problems.

RNA duplex is far easier to form than DNA, and it requires high energy to unzip the strands due to the accumulated effect of hydrogen bonding. Its high melting point is a major issue in this subject; recent studies show that by the alter of some 3'-5' nucleotides to the relatively active 2'-5' ones, or the introduction of under 17% inaccurate Watson-Crick base pairing, this unique property of RNA may be weakened. It has been shown that the usage of the entire 2'-5' nucleotide was still not attainable, as the RNA polymerase could only catalyst the reaction with a major amount of 3'-5' nucleotide. So that the complex of these two nucleotides has been stated.[13] On the other hand, the introduction of the inaccurate base pairing could effectively lower the stability of the RNA double helix, and in the actual base-pairing system, it is not paired with 100 per cent accuracy. However, only a certain amount of inaccuracy could be introduced to avoid mutation.

Another method states that the invade of some short oligonucleotides could synthesis the RNA strand displacement (Fig. 5). With the facilitating of activated 3'-amino nucleotides as an invader to let the displacement of RNA strand occur, the 5'-5'-imidazolium diribocytosine performed as the intermedia, we can copy the template of guanosine. Then the phosphate group on the imidazolium diribocytosine will be attacked by the hydroxyl group on the primer which let the primer extension occur and form a leaving group simultaneously. This reaction should be restricted in the circumstance with high strand concentration [18].

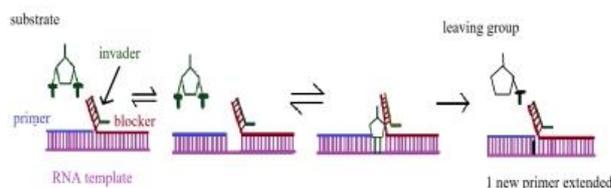


Figure 5. Primer extension by strand displacement.

The second problem was related to the catalysing effect of metal ions. Despite the role of metal ions on catalysing the RNA replication instead of protein, it still catalyses the degradation of the substrate as well as the RNA backbone by break the phosphodiester bond between the extended nucleotides, with the typical example being  $Mg^{2+}$  as we mentioned before. Also, the lipid-component selectively permeable membrane would not support the existence of  $Mg^{2+}$  [13]. The amount of  $Fe^{2+}$  should be inversely proportional to the  $P_{CO_2}$ (atmospheric pressure of  $CO_2$ ) [12]. But the  $CO_2$  will precipitate with  $Fe^{2+}$ , with the presence of UV radiation, the ferrous ions are oxidised to ferris ions. Also, under the condition of the primitive earth, the iron acts as electron donors before the accumulation of oxygen in the water.  $Fe^{2+}$  may easily form a complex with its surrounding constituents. So some people believe that the iron present in the primitive earth was in the form of iron oxide precipitation rather than iron cations. Most people accept the presence of cyanides in the origin of life, As the  $CN^-$  will strongly combine with  $Fe^{2+}$  and generate a complex  $Fe(CN)_6$ . The strong capability of iron to form complex seems obstacle the presence of  $Fe^{2+}$ .

The balance of pH with the efficiency requires at different stages in the RNA replication is another obvious issue, we decreased the PH value to adapt to the desired environment needed in the fork formation step to promote the likelihood for the reaction to occur. But in the strand renealing step, as the requirement of the neutral environment, we could not control the acidity in the optimum value for both stages.

Besides, as the error-prone primitive replication with cells, the base-bairing system might have inedequate fidelity, this causes the limitation in the length of RNA nucleotide elongation. However, the possibility of mutation creates competition among different phenotypes, so that due to natural selection, only the fittest phenotype will survive ultimately, and this provides the RNA with its best constitution and the most appropriate structure in evolution [19].

This paper suggests that the most severe issue that we have to tackle in this process is towards the dead-ended product generated during molecular replication. And perhaps the low fidelity of RNA copying is not such severe issue, as there is a wide range of reactions that do require the introduction of mismatching to promote the reaction pathway.

### III. CONCLUSION AND PROSPECTS

The informational molecular replication system is crucial in modern biology and the phase of the origin of life. The replication of nucleic acids, their strand

separation which is hard owing to the high  $T_m$  of RNA strand has been illustrated to happen with the introduction of mismatching, altering U to thio-U, which also contribute to the fidelity of RNA replication [20]. Their nucleotide revealing could occur in a high strand concentration atmosphere, to give high phenotypic effect to their host protocell. Also, in their model of strand displacement, it has been proposed to let it undergo the reaction with the assist of short oligonucleotides [21]. There may be some impediments generate as if all these steps should happen in succession, we have to concern about their efficiency and also, the environment we choose has to fit the demand of all those phases.

The non-enzymatic RNA replication do have some obstacles, scientists have suggested that the protocell's independence in its membrane's self-reproduction is a problem, such problem reveals in the intra-vesicular synthesis of the protocell membrane, which requires the enzyme involves in its genetic expression [22].

Tacitly, the significance is prosperous as if the RNA replication may happen in an enzyme-free condition. Then the RNA world hypothesis may be verified; on the other hand, this will imply in the protocells, replications seems unnecessary for the first heritable functional RNA as it will no longer need the help of protein synthesis. Despite those impediments, we can still find that the related topics about nucleic acid replication and the origin of life are profound in the later research. More people are eager to find ways to reemerge the protocell model. And the next step that scientists might consider is how to prove their hypothesis towards the origin of life realisable and figure out what kind of way the substances, including RNA and those molecules in the primitive earth exist and react, base on evidence that have been proved previously. And if the non-enzymatic RNA replication can be tested to be realisable, that way, it will be plausible to impose its method into model protocells, and in other words, try to mimic the origin of life.

#### CONFLICT OF INTEREST

The author declares no conflict of interest.

#### REFERENCES

- [1] W. K. Johnston, P. J. Unrau, M. S. Laurence, *et al.*, "RNA-catalyzed RNA polymerization: Accurate and general RNA-templated primer extension," *Science*, vol. 292, pp. 1319-1325, Dec. 2001.
- [2] J. W. Szostak, "The eightfold path to non-enzymatic RNA replication," *Journal of Systems Chemistry*, vol. 3, pp. 1759-2208, Feb. 2012.
- [3] J. D. Sutherland, "The origin of life—Out of the blue," *Angew. Chem. Int. Ed.*, vol. 55, pp. 104-121, 2016.
- [4] W. Zhang, T. Walton, L. Li, and J. W. Szostak, "Crystallographic observation of nonenzymatic RNA primer extension," *eLife*, vol. 7, pp. 1-2, May. 2018.
- [5] L. Zhou, S. C. Kim, J. W. Szostak, *et al.*, "Non-enzymatic primer extension with strand displacement," *eLife*, vol. 10, p. 770578, Nov. 2019.
- [6] Y. Qiao, M. Li, R. Booth., *et al.*, "Predatory behavior in synthetic protocell communities," *Nature Chemistry*, vol. 9, p. 110, 2017.
- [7] K. Ruiz-Mirazo, C. Briones, and A. Escosura, "New perspectives for the origins of life," *Chem. Rev.*, vol. 114, pp. 285-366, Oct. 2013.
- [8] T. Walton, W. Zhang, L. Li, *et al.*, "The mechanism of nonenzymatic template copying with imidazole-activated nucleotides," *Angewandte Chemie*, vol. 131, pp. 10926-10933, March 2019.
- [9] J. Lieberman, "Tapping the RNA world for therapeutics," *Nat. Struct. Mol. Biol.*, vol. 25, p. 357, 2018.
- [10] S. Roland and A. Pyle, "Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry," *Chem. Rev.*, vol. 107, pp. 97-113, Dec. 2007.
- [11] N. Ebner and S. Haehling, "Iron deficiency and heart failure: A practical guide," *Nutrients*, vol. 5, pp. 3730-3739, Sept. 2013.
- [12] S. Sadjjadi and J. Massoud, "Study on the latent iron deficiency in patients with Trichuriasis," *Iranian Journal of public health*, vol. 27, pp. 37-46, 1998.
- [13] W. Jahnen-Dechent and M. Ketteler, "Magnesium basics," *Clinical Kidney Journal*, vol. 5, pp. 3-14, Feb. 2012.
- [14] B. L. Mordike and T. Ebert, "Magnesium: Properties-application-potential," *Materials Science and Engineering*, vol. 302, pp. 37-45, 2001.
- [15] F. Albarede, F. Thibon, J. Blichert-Toft, *et al.*, "Chemical archeoceanography," *Chem. Geo.*, vol. 548, p. 119625, 2020.
- [16] N. Tosca, *et al.*, "Products of the iron cycle on the early Earth," *Free Radical Biology and Medicine*, vol. 140, pp. 138-153, Aug. 2019.
- [17] C. Bonfio, E. Godino, M. Corsini, *et al.*, "Prebiotic iron-sulfur peptide catalysts generate a pH gradient across model membranes of late protocells," *Nature Catalysis*, vol. 1, pp. 616-623, July. 2018.
- [18] Y. Ohara, M. K. Readan, K. Fujikura, *et al.*, "A serpentinite-hosted ecosystem in the southern mariana forearc," *Ecology*, vol. 109, pp. 2831-2835, Feb. 2012.
- [19] L. M. R. Keil, *et al.*, "Proton gradients and pH oscillations emerge from heat flow at the microscale," *Nature Communications*, vol. 8, pp. 1-9, Dec. 2017.
- [20] N. Kitadai and S. Maruyama, "Origin of building block of life: A review," *Geosci. Frontiers*, vol. 9, pp. 1117-1153, 2018.
- [21] J. W. Szostak, "The narrow road to the deep past: In search of the chemistry of the origin of life," *Angew. Chem. Int. Ed.*, vol. 56, pp. 11037-11043, 2017.
- [22] A. Kun, M. Santos, and E. Szathmary, "Real ribozymes suggest a relaxed error threshold," *Nature Genetics*, vol. 37, pp. 1008-1011, Sept. 2005.

Copyright © 2021 by the authors. This is an open access article distributed under the Creative Commons Attribution License ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

**Xinyu Wu** was born on April 6<sup>th</sup> 2004. She is currently is a student studying Alevel courses in Beijing Jingshan School International Department of Tangshan, China. She is interested in biochemistry and chemistry, especially the research towards the origin of life and molecular replication. She won silver medal in UK Chemistry Olympiad in 2020 and is planning to study in Britain in the future to exploit biochemistry in depth.