Role of metal ions and their complexes in the catalysis of pertinent prebiotic reactions

A Thesis

Submitted in partial fulfillment of the requirements

Of the degree

Doctor of Philosophy

By

Shikha Dagar

20152012

Indian Institution of Science Education and Research (IISER) Pune

DEDICATION

I dedicate this work to my parents

Kuldeep Singh Dagar and Roshani Devi

Certificate

Certified that the work incorporated in the thesis entitled "Role of metal ions and their complexes in the catalysis of pertinent prebiotic reactions", submitted by Ms. Shikha Dagar, was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Dr. Sudha Rajamani

15.07.2022

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources that have not been properly cited or from whom proper permission has not been taken when needed.

Shikha Dagar

15.07.2022

Acknowledgments

I would like to thank my thesis supervisor Dr. Sudha Rajamani for her guidance and unconditional support through my doctoral research. I am grateful beyond words to her, for her continuous encouragement throughout. I am indebted to her for providing me with opportunities to develop strong research skillset and for the necessary training and resources required for completing this thesis.

I would like to thank other members of my research advisory committee; Prof. Nandita Madhavan (IIT Bombay), Prof. V.G. Anand and Dr. Jeetender Chugh for their valuable inputs in my research work. I am grateful for their time and constant encouragement throughout my doctoral studies. I would like to thank extend my special thanks to Prof. V.G. Anand and his lab members for their help and useful discussions.

I would like to thank Dr. Mayurika Lahiri and Dr. Gayathri Pananghat for providing me with the opportunity to work as rotation student in their laboratory and training.

I would also like to express my gratitude to all the present and past COoL lab members for their support and for making the work environment lively and fun. I thank members of the Mass Spectrometry facility and Microscopy facility in the Department of Biology and HRMS facility, FE-SEM facility at IISER-Pune, for their useful training, cooperation and help in performing the experiments. I also thank amicable non-teaching staff of the Department of Biology for their support and easy disposal of the formalities necessary to drive my academics.

I am really grateful to my friends, especially Susovan, Sasank, Vinayak, Sutirtha, Sandhya, Mishika, Joyeeta, Rupali, Mrinmayee, Shruti, Angira, Avisikta and Prachi for their support, help and making my journey in IISER Pune, a memorable experience. I extend my deepest gratitude to my family, my parents, my brothers, my sister and my "brother" Susovan, for their unparalleled love, support and encouragement to achieve my dreams. I am grateful to my family and friends for keeping faith in me and showing confidence in me.

Table of Content

List of Figures

List of Tables

Chapter 1

Introduction

1.1. Introductions

The extant complex life that is pervading is the outcome of evolution that happened over billions of years. Earth is hypothesized to have been formed around ~4.6 billion years ago. This is thought to have been followed by the formation of a stable hydrosphere around ~4.2 billion years ago (Hadean era), where interesting prebiotic chemistry could have been facilitated[1]. Some of these prebiotically pertinent reactions would have led to the formation of primitive forms of biologically relevant molecules on the early Earth[2]. These would have been mainly nonenzymatic, governed and driven by environmental factors, and influenced by co-occurring molecules in the heterogenous prebiotic soup. Scientists working in the area of chemical origins of life continue to ponder how and where life might have originated. A related aspect that is debated often in the field is the features of life and how one can actually define it. Despite competing theories, one definition that was put forward by Dr. Gerald Joyce, and adopted by NASA, defines life in the following way: Life is "A self-sustaining chemical system capable of Darwinian evolution"[3]. Alternatively, life can be described by the fundamental hallmarks that define it, which are: information encoding and replication, being metabolically active, and maintaining its integrity by invoking relevant boundary conditions[4]. Given the complexity intrinsic to each of these hallmarks, the spontaneous emergence of a primitive life form from a mix of chemical entities seems improbable. Nonetheless, on a closer look, it becomes obvious that this would have been rendered possible by the formation of complex chemical molecules, by a concerted action of the geological, physical and chemical variables/constraints that were present at various times in the early history of life on Earth. Many of these complex molecules would have evolved to result in emergent properties, kickstarting a process that eventually led to the chemical origins of life.

In extant life, biomolecules are utilized to efficiently perform various functions that are crucial for it to sustain an out-of-equilibrium state. Nucleic acids (DNA or RNA) predominantly store and process genetic information, while proteins perform the

metabolic functions and lipid bilayers provide a boundary condition that separates the encapsulated system from the environment. Along with providing integrity to the system, these membrane compartments also control the exchange of components with the outer environment[5]. The evolution of such efficient and specific biomolecules to perform the varied functions of life, took billions of years to evolve. However, at the earlier stages of life's emergence, these functions might have been carried out by simpler, prebiotically plausible molecules. These could have included oligopeptides/primitive catalysts, a primitive form of an informational encoding system (e.g., simpler nucleic acids) and relatively simple amphiphilic compartments. In this regard, several hypotheses have been put forward that signify the central role of one component over the other. These include the "Lipid World hypothesis", the "RNA World hypothesis" and the "Metabolism first hypothesis"[6–8].

The Lipid World hypothesis emphasizes the crucial role of compartments in maintaining the integrity of the system[9]. On the other hand, the metabolism first hypothesis highlights the essentiality of a primitive form of metabolism in order to maintain the system at a far-from-equilibrium state[10–12]. The most widely studied RNA World hypothesis, argues that the RNA molecule, owing to its capability to act as both a genetic material as well as a catalyst, would have possibly been the first biomolecule of consequence to have emerged on the early Earth. Although insightful, all these aforementioned features would have been crucial for the emergence of a self-sustaining primitive cell (protocell). And, this would have been facilitated by the heterogeneity of the prebiotic soup that would have promoted cross-talk between pertinent prebiotically plausible molecules/co-solutes performing these key functions thus aiding the emergence of a protocell. In this chapter, we review relevant studies pertaining to different hypotheses and attempt to bridge these studies for a more holistic understanding of the prebiotic context. We specifically discuss how accounting for the heterogeneity of prebiotic soup, can indeed provide important insights towards understanding some of these aspects.

1.2. Compartments

One of the salient features mandatory for a system to be called 'alive' is compartmentalization, which allows for its maintenance and sustenance at a far-from-equilibrium state[5,7,9]. Similar to contemporary cell compartments, primitive compartments would have allowed for the exchange of solute molecules with the bulk environment along with maintaining the integrity of a system. Towards this, several membrane-less systems such as aqueous two-phase systems[13], coacervates[14], and membrane-bound compartments[15,16], have been explored. As for the membrane-less moieties, biphasic systems are the major contenders and these are considered to have preceded membrane-bound organelles. Conversely, membrane-bound structures result in a triphasic system with an aqueous bulk solution outside, the hydrophobic boundary in between and an aqueous lumen in the interior, similar to what is seen in the extant membrane-bound cell. Such triphasic systems have the ability to encapsulate solute molecules within and thus, act as a selectivity barrier between the outside environment and the encapsulated system[16]. Contemporary cell membranes utilize phospholipid-based entities that primarily are diacyl chains linked to a glycerol moiety with a phosphate head group. The phosphate head group is further derivatized to form different head groups such as choline group in phosphatidylcholine, serine in phosphatidylserine, to name a few. This derivatization of the phosphate head group further adds to the complexity of the phospholipid diacyl structure. Additionally, the low yields of their prebiotically plausible synthesis have rendered their prebiotic relevance debatable[17,18]. Given this, protocells are thought to have been composed of simpler, prebiotically relevant single-chain amphiphiles (SCAs)[5,19,20]. Herein, we discuss some of the model protocellular membrane systems that have been explored in the field.

1.2.1. Membrane-bound compartments

Extant cell membranes are complex structures that are impermeable to polar solute molecules[5]. To overcome this, cells utilize sophisticated protein machinery that facilitates the exchange of solute molecules selectively through it. However,

the spontaneous emergence of such proteins on the early Earth is improbable. Therefore, the questionable prebiotic abundance of phospholipids, along with the limited permeability of these membranes, renders them unsuitable as protocellular membranes. Because of the aforementioned issues, several other systems composed of structurally simpler SCAs have been explored and characterized.

Towards this, systems composed of fatty acids (FAs) have been widely explored[21,22]. FAs can be synthesized by abiotic Fisher-Tropsch- Type (FTT) synthesis, and have also been detected in several meteoritic samples[23]. FAs with the chain length >C8 are known to possess high CBC (critical bilayer concentration; i.e., the minimum threshold concentration which is required to self-assemble into a bilayer). When compared to phospholipids, FA vesicles are dynamic in nature and this makes them relatively more permeable. However, FAs can self-assemble only around the pH near the pKa of their head group, which limits their vesicle formation to a narrow pH regime[21,24]. Moreover, the self-assembled structures are sensitive to the surrounding temperature as well as the presence of metal ions[5,25]. This would have posed a barrier for FA-based vesicles to act as protocellular membranes as prebiotically relevant reactions are hypothesized to occur at a wide range of pH and temperature[26–29]. Additionally, metal ions have been demonstrated to be crucial for a range of reactions, including folding and functioning of ribozymes, in primitive metabolic reactions like reverse tricarboxylic acid (rTCA) cycle, gluconeogenesis etc.[30–33] This necessitates the requirement of a robust prebiotically relevant protocellular membrane systems.

In this context, a binary system composed of fatty alcohols along with FAs, was demonstrated to self-assemble into vesicular structures at lower CBC that were also stable in alkaline regimes[25]. Subsequently, a tertiary system containing fatty alcohols and glycerol derivative of fatty acids, along with FAs, was shown to enhance the thermostability of self-assembled structures. Later on, the self-

assembly behaviour of tertiary systems under varying selection conditions was investigated. Similar to the previous results, the addition of fatty alcohol was observed to stabilize the FA-based membranes under alkaline conditions. The addition of glycerol monoester derivatives of FAs, was observed to stabilize it against metal-ion induced flocculation. When multiple selection conditions (dilution, presence of salt and alkaline pH) were applied, the compositionally diverse systems i.e., tertiary systems, were seen to survive when compared to the pure fatty acid or binary systems composed of either glycerol monoester or fatty alcohol, along with the fatty acid. Although insightful, even tertiary systems composed of fatty acid, with its fatty alcohol and glycerol monoester derivatives as co-surfactants, could not facilitate self-assembly in the acidic regimes. Several prebiotically pertinent polymer formation reactions such as RNA oligomerization and peptide formation are shown to occur at acidic pH and high temperature[28]. Thus, the sensitivity of FA-based vesicles to acidic pH poses a significant challenge for these systems to effectively act as protocellular membranes. In this regard, other systems have been investigated for their ability to form a bilayer in the acidic pH. The admixing of sodium dodecylbenzene sulfonate (SDBS) with decanoic acid vesicles was observed to form vesicles even at acidic pH (up to pH 4.3)[34]. In a study where another SCA with a cyclic phosphate head group (cyclophospholipid) was characterized, the authors observed formation vesicles in a wide range of pH (4-10)[35]. In very recent work from our lab, systems consisting of single-chain alkyl phosphate (i.e., dodecylphosphate; DDP), in isolation or in conjunction with fatty alcohols (in a binary mixture), was demonstrated to form pH-responsive membranes all the way from pH 2 to pH 10[36]. DDP, when used in conjunction with phospholipid, was also shown to modulate the properties of usually not-very-dynamic phospholipid membranes. Such tuneability would have been beneficial towards constructing protocellular membranes with several desired properties on the early Earth.

1.2.2. Protocellular model membranes vs Contemporary cell membranes

As discussed earlier, several systems containing one kind of amphiphile or a mixture of different amphiphile (heterogeneous) systems have been explored as protocellular membranes. Predominant studies in this regard have used FA-based vesicles as model membranes. Although insightful, how the transition from these model protocellular systems to the extant phospholipid-based cell membranes would have happened is still unclear. Some studies have suggested hybrid/blended membranes composed of both fatty acids and phospholipids to have been an intermediate step in this transition process[37]. The direct transition to the aforementioned kind of blended membranes in a composite prebiotic soup, which is very heterogeneous where several other amphiphiles would have also been present, does not seem straightforward. Therefore, it is logical to assume that the composition of protocellular membranes during the various stages of this transition would have been governed by environmental conditions and could have involved several intermediate stages, eventually resulting in the evolution of extant membranes[5]. These latter membranes that are composed mainly of phospholipids, would have then incorporated proteins and other types of amphiphiles (cholesterol, ceramide etc.) in order to modulate their properties and promote communication with the external environment; all of which is essential for facilitating cell physiology in an effective manner.

1.3. Minimal genetic information

Another key feature of a living system is the presence of genetic material[38]. This characteristic is crucial for the storage and propagation of genetic information from one generation to another, thus, imparting the capability to initiate the process that eventually leads to Darwinian evolution. Contemporary biology utilizes nucleic acids, specifically, RNA and/or DNA for this purpose. Although both DNA and RNA are capable of storing genetic information, the higher reactivity of RNA due to the presence of a hydroxyl group at the 2′ position allows it to fold upon itself, in turn aiding its capability to act as a catalyst in addition to

being a genetic material[39]. This concept is explored widely as the 'RNA World Hypothesis' and is discussed below and in chapters 2, 3 and 4 in detail.

1.3.1. RNA World Hypothesis

This well-explored hypothesis was first proposed in the 1960s due to the capability of RNA to fold upon itself and attain secondary structures, thus can act as a catalyst along with being a genetic material[39]. It was later supported by the discovery of ribozymes (catalytic RNAs) in the ribosome, which facilitates the translation of RNA into proteins. Although rational, the RNA World hypothesis presumes the presence of reasonably long stretches of RNA that could have attained interesting structures to let some of them to act as catalysts. The process of formation of RNA on the prebiotic Earth is itself non-trivial and would have been predominantly driven by nonenzymatic processes because the presence of complex protein machinery on the early Earth would have been implausible. Hitherto, scientists have demonstrated the formation of RNA constituents i.e., nucleosides and their phosphorylation to form nucleotides, which is thought to be catalyzed by prebiotically relevant metal ions and minerals[40,41]. These nucleotides are required to undergo thermodynamically uphill condensation reactions to form functional stretches of RNA[26,27,42]. Towards this, several studies have attempted to demonstrate the enzyme-free oligomerization and template-directed information propagation of these molecules as discussed below.

1.3.1.1. Nonenzymatic oligomerization of nucleotides

The spontaneous condensation of nucleotides to form RNA strands involves the removal of a water molecule in the process, and would have been thermodynamically unfavourable in an aqueous prebiotic pool. Majority of studies that attempted to demonstrate enzyme-free polymerization in this context, typically used chemically modified nucleotides (refs). They also evaluated the catalyzing effect of cations present in minerals like montmorillonite, on the nonenzymatic oligomerization of nucleotides[43]. These studies showed that some

metal ions (Li$^+$ and Na$^+$) were better catalysts than others (K$^+$, Cu^{2+}, Ni^{2+} and Mg^{2+} [44,45]. These chemically activated nucleotides possess a good leaving group attached to their 5′ phosphate moiety; for e.g., imidazole activated-nucleoside-5′-monophosphates[46]. The presence of a good leaving group facilitates the formation of a phosphodiester bond between two ribonucleotides when there is a nucleophilic attack by the 2′ or 3′ hydroxyl group of an incoming nucleotide, as shown in Figure 1.1.

Although the prebiotically plausible synthesis of such activated nucleotides has been demonstrated, their availability in significant amounts on early Earth is still uncertain[47]. Moreover, the unstable nature of these activated nucleotides toward high temperatures and aqueous conditions, makes their potential role as substrates for enzyme-free oligomerization highly debatable.
Towards this, the nonenzymatic oligomerization of non-activated nucleotides (nucleoside 5′-monophosphates; 5′-NMPs) has been investigated under terrestrial volcanic geothermal conditions[26–28]. Such niches that are characterized by high temperatures and highly acidic pH (pH 2), were observed to be conducive for oligomerization to result in RNA-like polymers. The systematic characterization of these polymers showed that they contained abasic sites with just one nucleobase seemingly attached to the sugar-phosphate backbone due to the depurination which is facilitated under these conditions (Figure 1.2)[28]. This inspired us to investigate other intrinsically activated nucleotides such as cyclic nucleotides (cNMP) and nucleoside triphosphates (NTPs) for their potential to oligomerize yielding RNA polymers. This is discussed in detail in chapters 2 and 4.
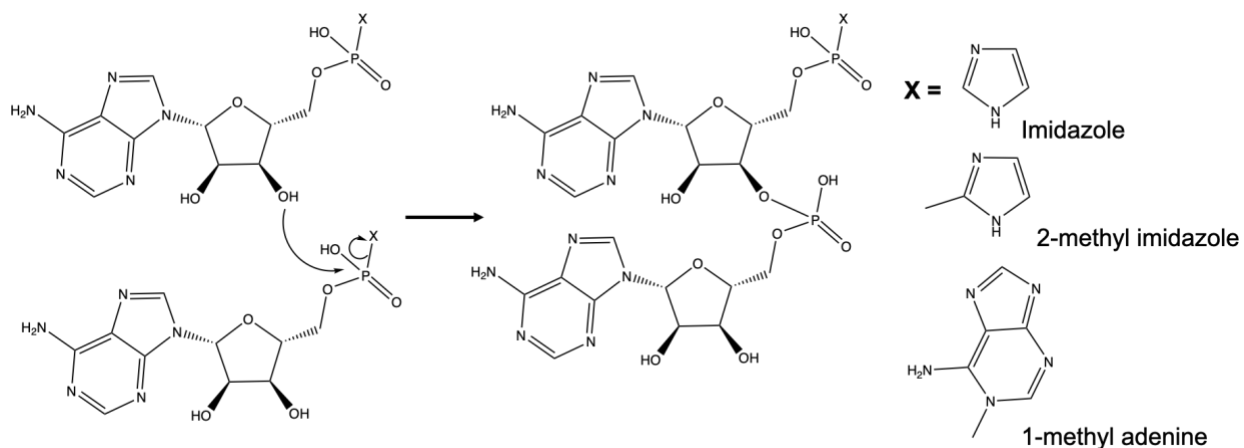
Figure 1.1: Mechanism depicting the RNA oligomerization reaction. X is a good leaving group and acts as a sink for electrons, facilitating the attack of 2′ or 3′-OH on the phosphate group thereby helping in phosphodiester bond formation (for simplicity only 2′ attack is shown). Structures drawn using ChemDraw 20.0
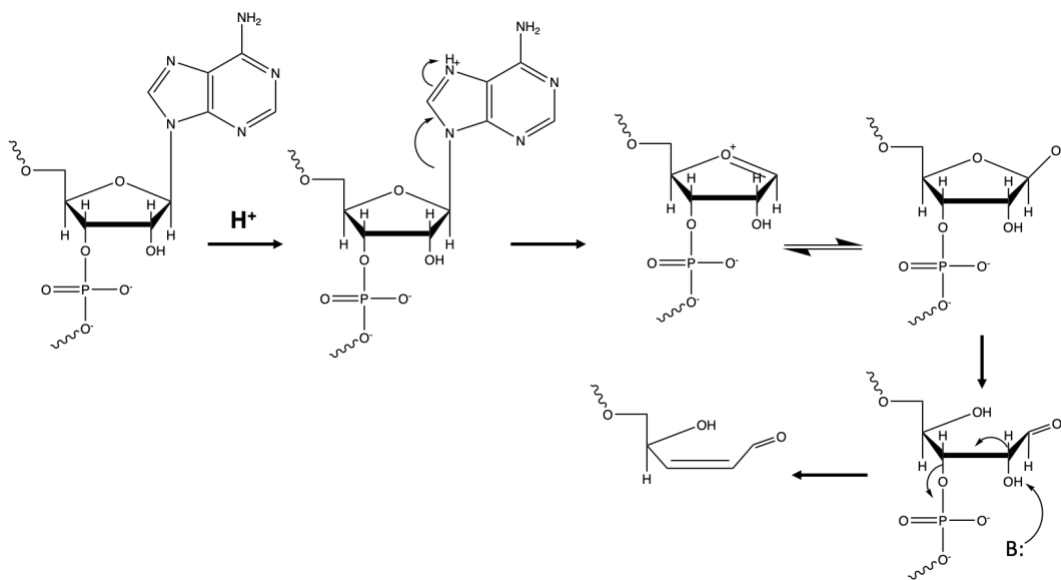
Figure 1.2: Mechanism of acid-catalyzed depurination and cleavage in RNA shown for adenine base (as an example). This involves protonation at N-7 of adenine, followed by breakage of the N-glycosidic bond resulting in depurination. The depurinated product can further undergo ring opening and elimination to form an open-chain aldehyde. Structures are drawn using ChemDraw 20.0

1.3.1.2. Nonenzymatic template-directed replication of RNA

Once the genetic information (in this case RNA) is formed, the next logical step to discern is their faithful replication to generate more copies of themselves. The

replication/ propagation would have been enzyme-free owing to the improbability of enzymes on the early Earth and would also have been driven by molecular and physicochemical environmental conditions[48]. Three current models for the nonenzymatic replication[49] (Figure 1.3) are as follows: a) First one pertains to base-pairing of monomeric nucleotides against a preformed template, facilitating their clustering and eventually resulting in a phosphodiester bond between neighbouring nucleotides to yield a daughter strand. b) The second model in this regard involves step-wise addition of nucleotides to a small oligomer annealed to the preformed template that acts as a primer and forms a template-primer duplex. c) The third model is inspired by the previous two models and suggests annealing of multiple small oligomers to the same template based on base complementarity. These might result in gaps where the insertion of monomeric nucleotides and their subsequent ligation results in a fully formed daughter strand.
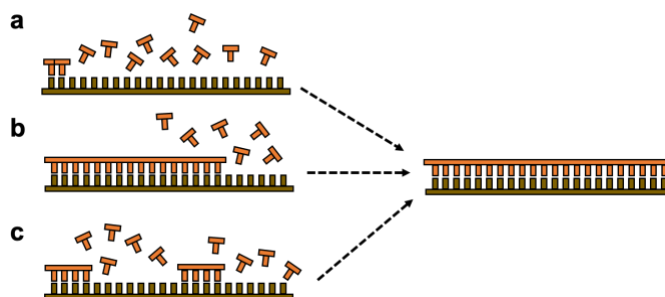
Figure 1.3: Models depicting nonenzymatic template-directed replication to yield a fully replicated daughter strand. A) by base-pairing of monomeric nucleotides on an RNA template. B) step-wise addition of nucleotides on a primer annealed to a template. c) annealing of small oligomers on the template, followed by nucleotide insertion in the gaps and their subsequent ligation. Adapted from Kaddour, H. & Sahai, N. Life **4**, 598–620 (2014).

A substantial amount of work has been done in this endeavour. Similar to polymerization studies, most of the studies in the aforesaid regard have utilized 'activated nucleotides' where the 5′- phosphate group is activated with a good leaving group, predominantly imidazole or imidazolide derivates[50–53]. In addition to facilitating the primer extension by overcoming the thermodynamic barrier for

bond formation, this activation of incoming nucleotides also allows for the monitoring of these reactions within laboratory time scales. The detailed mechanistic characterization of reactions using imidazole-activated nucleotides has demonstrated the bond formation by forming a 5′-5′-imidazolium bridged dinucleotide intermediate[54]. Other activating groups like carbodiimide or 1-hydroxy-7-azabenzotriazole (HOAt) have also been shown to facilitate nonenzymatic template-directed replication[55]. One important aspect of replication is its fidelity as the addition of a mismatch nucleobase can lead to the loss of functional information. In 2010, Rajamani et al. have shown that the rates of subsequent nucleotide addition decrease by more than 2 orders of magnitude upon the addition of an incorrect nucleotide[50]. This decrease in the rate results in the stalling of replication after a mismatch addition, thus providing a proof of principle for faithful copying of the genetic information even in the absence of enzymes.

A subsequent related study demonstrated that the presence of background molecules (co-solutes) such as lipids and/or PEG (as a proxy for molecular crowding) results in an increase in the frequency of mismatch pairs[51]. This was indeed interesting as the prebiotic soup is thought to be heterogeneous in nature and would have been compositionally diverse. Therefore, the presence of co-solutes in such scenarios would have allowed for the exploration of different sequences and thereby different functions, over reasonable timescales. Nonetheless, the question of the ready availability of a substantial pool of such intrinsically activated nucleotides led researchers in a parallel study to use non-activated 5′-NMP. They studied acid-catalyzed template-directed primer extension using 5′-NMP but the extended primer was observed to contain abasic site[48]. Similar to the oligomerization reactions, this was due to the cleavage of glycosidic bonds under harsh conditions such as acidic pH (2) and high temperature (90°C). This further emphasized the necessity and importance of the use of intrinsically activated prebiotically relevant nucleotides, which do not require harsh conditions to oligomerize and propagate information readily under

early Earth conditions. The ability of intrinsically activated prebiotically relevant nucleotides such as cNMP to propagate template-directed information is discussed in details in chapter 3.

1.3.1.3.　Ribozyme-mediated catalysis and role of metal ions

The emergence of a polymerase and self-replicase ribozymes lies at the heart of the 'RNA World Hypothesis'. Under early Earth-like conditions, these and other ribozymes are thought to have facilitated several prebiotically pertinent reactions that eventually resulted in protocell formation. For e.g., these ribozymes would have facilitated the formation as well as propagation of RNA, a crucial feature for the emergence and sustenance of protocells[56]. In 1993, a pioneering study demonstrated the emergence of a ligase ribozyme when a random pool containing 10^{15} molecules of RNA was subjected to repetitive in-vitro selection cycles. This prompted a series of studies investigating the catalytic role of RNA, their in-vitro emergence and evolution using Systematic Evolution of Ligands by Exponential Enrichment (SELEX), to produce different ribozyme sequences that resulted in different functions[56–58]. Ribozymes possessing replicase, ligase and polymerase activity have been developed by various groups through in-vitro evolution. Although insightful, the spontaneous emergence of ribozymes which are itself long stretches of structured RNA, is non-trivial in the absence of specialized protein machinery. Another salient feature for most of these ribozymes is the requirement of appropriate amounts of metal ions to maintain their structure and function[59–62]. RNA phosphodiester backbone possesses a negative charge that results in repulsion and, in turn, hindering its folding to result in secondary structure formation. Metal ions are known to coordinate with the negatively charged phosphate groups of the RNA backbone, thus reducing this repulsion and facilitating folding that is central to its function. In addition to this, metal ions are also observed in the active site of several ribozymes where they directly take part in the catalytic processes. Most of the ribozymes rely on the presence of Mg^{2+} ion that is coordinated in its active site for their functioning[63]. Previous studies have also evaluated a significant increase in catalytic activity by

substituting Mg^{2+} with Fe^{2+} in two different ribozymes, i.e., L1 ribozyme ligase and hammerhead (a self-cleaving) ribozyme[64]. This further emphasize the crucial role of metal ions in the structural stability and functioning of ribozymes, which are key for a putative RNA World. Even though these SELEX based studies have been truly revealing, the iterative in-vitro selection is generally carried out under controlled reaction conditions, which would have not been the case on early Earth. This, therefore, leaves a gap for discerning realistic approaches that would result in the nonenzymatic emergence of fully functional ribozymes.

1.4. Minimal metabolism

Lastly, for the sustenance of any living system, there is the requirement of access to (minimal) metabolic pathways(s) that can take care of the energy requirements. In this regard, the two unresolved questions in the context of early life are as follows: 1) What kind of reactions are crucial for the sustenance of primitive cellular life? 2) Which prebiotically plausible catalysts would have aided various steps of these crucial reactions? On the early Earth, metabolism is hypothesized to have been dependent mainly on redox reactions, as is seen in carbon fixation, sulfur reduction or nitrogen fixation, wherein energy is produced via electron transfer[65,66]. Prior to the onset of an oxygenic atmosphere, nitrogen and sulfur reduction are thought to have been the major driving forces for facilitating metabolism[67]. Using phylogenetic analysis, enzymes involved in the aforementioned reactions have been traced from higher organisms (like plants), all the way back to eubacteria, archaebacteria, and cyanobacteria[68]. Additionally, some of the core metabolic cycles such as the (reverse) tricarboxylic acid cycle, pentose phosphate pathway, glycolysis and gluconeogenesis etc.[13,32,69], are also considered as ancient due to their favourable thermodynamics and universality.

Previous literature has investigated the role of metal ions in primitive metabolic networks, including the reverse tricarboxylic acid (rTCA) cycle, glycolysis, pentose phosphate pathway etc.[31,32,69,70], that underscore the central role of

simple inorganic molecules in prebiotic catalysis. Importantly, about one-third of contemporary enzymes are metalloenzymes, i.e., they depend upon metal ions for their functioning[71]. This emphasizes the crucial role that metal ions would have played, in the emergence and early evolution of such metalloenzymes. This has inspired researchers to explore primitive metabolism and to specifically investigate the capability of metal ions/ metal-inorganic scaffolds (as present in minerals) in catalyzing several prebiotic pertinent reactions. In this regard, Zn^{2+}, Cr^{3+} and Fe^0 were shown to promote multiple reactions of the rTCA cycle[31] in an acidic aqueous environment. Ralser's group has demonstrated the nonenzymatic formation and interconversion of metabolites involved in the pentose phosphate pathway, with the help of a mimetic Archean ocean that is rich in ferrous (Fe^{2+}) ions[69,70]. Moran's group has also demonstrated the catalytic role of native transition metals (Fe^0, Ni^0 and Co^0) in CO_2 reduction, to yield the intermediates (such as acetate and pyruvate) of the Wood-Ljungdahl pathway[30].

The iron catalysis aspect pertaining to life's origins, has been well explored as Iron-sulfur World hypothesis, which was proposed by Gunter Wachterhauser[72–74]. It hypothesizes that life arose as a geochemical process from inorganic starting materials on the surface of sulfide minerals (such as pyrite), in the vicinity of deep-sea hot springs. Based on all the aforementioned aspects, it seems reasonable to argue that the first cellular life on Earth might have utilized relatively simpler protoenzymes and/or a more basic, related scaffold to carry out metabolic redox reactions. The capability of pyrite to fix nitrogen under UV irradiation has been demonstrated further supporting the hypothesis[75]. Minerals such as CdS (greenockite), ZnO (Zincite), CdSe and FeS_2 (pyrite) have recently been shown to catalyze transmembrane electron transfer reactions (TMETR) upon irradiation[76]. Although insightful, these studies do not bridge the evolutionary pathway of metalloenzymes from a geochemistry-based perspective.

Iron-sulfur cluster proteins and metalloporphyrins consisting of a metallic core with a porphyrin ring and is surrounded by a protein scaffold, are some of the most ancient of enzymes present universally in the tree of life[68,77] (Figure 1.4). A pioneering study by Melvin Calvin in 1959 showed that the efficiency of catalyzing the hydrolysis of peroxide increases by up to 10^3 times when ferric ion is complexed with porphyrin heterocycle to form ferric-porphyrin coordinated complex[78]. The activity was found to be 10^7 times higher than just the ferric ion in a native peroxidase[78]. This emphasized that the addition of, and increase in the complexity of a scaffold, increases the reaction efficiency and its rate. With this line of thought, some previous studies have demonstrated the formation of iron-sulfur clusters under prebiotically plausible conditions and their capability to accept and donate electrons[79–82]. Subsequently, a potential pathway from iron-sulfur catalysts to iron-sulfur cofactor containing peptide catalysts was also recently demonstrated[83]. These iron-sulfur peptides were able to generate a pH gradient across the membrane, which is a salient feature of contemporary life. These studies have indeed added to our understanding of the emergence of metalloenzymes on the early Earth, mainly the ones that belong to the iron-sulfur class of proteins. However, studies pertaining to the emergence and early evolution of other ancient metalloenzymes such as plastocyanin, cytochrome c and symerythrin, are largely unexplored. Thus, relevant studies on iron-sulfur clusters and a few studies towards metalloporphyrins, set the stage for further exploration of these minimal enzymes as well as other small folds pertinent to functional aspects. Inspired by this in chapter 5, we investigated the capability of different metalloporphyrins to catalyze proxy essential redox reactions.
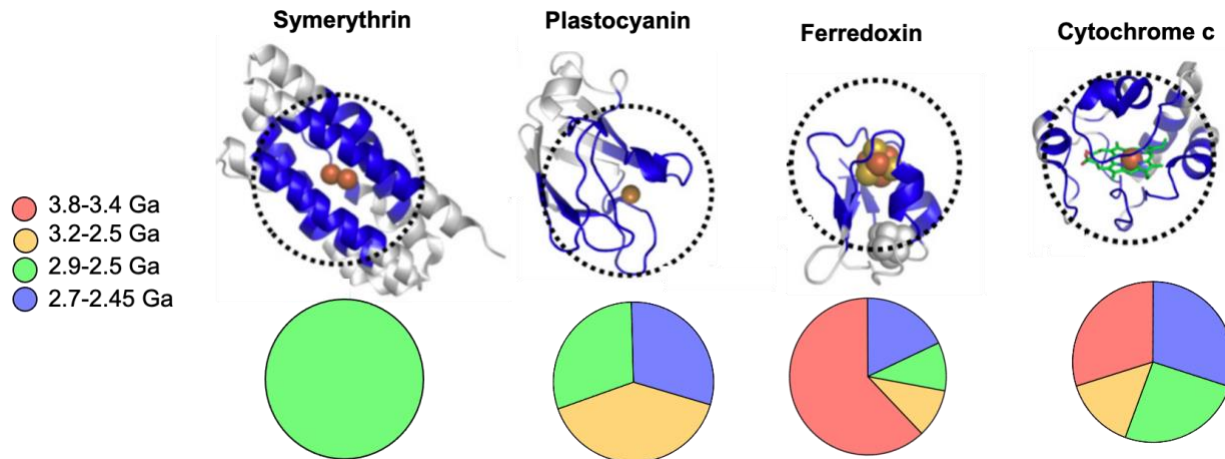
Figure 1.4: Estimated age of modules relative to the four stages of global electron transfer network expansion, inferred from geological records. Adapted from [84]: Raanan, H. et. al, PNAS, **115**, 1280–1285 (2018).

1.5. Discussion

The first forms of cellular life (protocells) on the early Earth are posited to be "a self-sustaining chemical system capable of Darwinian evolution"; composed of primitive genetic polymers and metabolic reaction networks encapsulated in a membranous compartment[7]. Given this, we first set out to investigate the formation and replication of RNA using prebiotically pertinent nucleotides. RNA is thought to have been one of the first biomolecules to have emerged on the early Earth due to its intrinsic ability to both store and propagate information, and to facilitate catalysis of reactions. Previous studies have investigated the nonenzymatic oligomerization/replication of chemically activated nucleotides whose presence on the prebiotic Earth in significant amounts is debatable. Therefore, we evaluated the oligomerization using intrinsically-activated yet prebiotically relevant nucleotides i.e., cyclic nucleotides (cNMPs) and nucleoside 5′-triphosphates (NTPs). Both, cNMPs and NTPs resulted in oligomers when subjected to repeated dry-wet cycles, a feature common on the early Earth just as it is today. Furthermore, we also tested the role of metal ions on these aforementioned processes by characterizing these reactions in hot spring water samples from Astrobiologically relevant sites in Ladakh, India[85]. Subsequently, we investigated the ability of cNMPs to propagate genetic information in enzyme-free template-directed

replication reactions. These reactions showed the addition of intact nucleotides to the primer, which resulted in the formation of informational polymers, something that has been a challenge when using potentially more readily available non-activated monomers. In all, these studies underscored the relevance of both cNMPs and NTPs to act as building blocks, in the formation and propagation of genetic information on the early Earth. In addition to the aforesaid, we were also keen on investigating transition metal ion-mediated oxidation reactions, specifically in the context of primitive metabolism. Few studies in the past showed that free metal ions and select metal-complexes can catalyze certain prebiotically relevant reactions. Nonetheless, the transition from metal ions to extant metalloenzymes has continued to largely remain unclear. To unravel the plausible route, we evaluated the influence of a pre-biologically and biologically relevant porphyrin scaffold, on metal ion-mediated oxidation reactions. The presence of porphyrin was observed to modulate the catalytic activity of certain metal ions, highlighting the selective advantage that these metal ions would have had in the presence of scaffolds like porphyrin. Altogether, our findings have helped to delineate aspects pertaining to the emergence and propagation of primitive genetic polymers. Furthermore, we have also been able to ascertain a framework for the evolution of minimal enzymes, thereby furthering the field's insights on processes central to the emergence of early life.

1.6. **References**

1. Sasselov, D. D., Grotzinger, J. P. & Sutherland, J. D. The Origin of Life as a Planetary Phenomenon. 1–13 (2020).
2. Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214–221 (2002).
3. Gánti, T. *The Principle of Life*. (Oxford University Press, 2003).
4. Ruiz-Mirazo, K., Peretó, J., The, A. M.-O. of L. and E. of & 2004, U. A universal definition of life: autonomy and open-ended evolution. *Orig. Life Evol. Biosph.* **34**, 323–346 (2004).
5. Sarkar, S. *et al.* Prebiological Membranes and Their Role in the Emergence of Early Cellular Life. *J. Membr. Biol.* **253**, 589–608 (2020).

6.   Lanier, K. A. & Williams, L. D. The Origin of Life: Models and Data. *J. Mol. Evol.* **84**, 85–92 (2017).

7.   Koch, A. L. & Silver, S. *The first cell. Advances in Microbial Physiology* **50**, (2005).

8.   Otto, S. An Approach to the De Novo Synthesis of Life. **13**, 54 (2022).

9.   Segré, D., Ben-Eli, D., Deamer, D. W. & Lancet, D. The Lipid World. *Orig. Life Evol. Biosph.* **31**, 119–145 (2001).

10.   Braakman, R. & Smith, E. The compositional and evolutionary logic of metabolism. *Phys. Biol.* **10**, (2013).

11.   Sojo, V., Herschy, B., Whicher, A., Camprubi, E. & Lane, N. The origin of life in alkaline hydrothermal vents. *Astrobiology* **16**, 181–197 (2016).

12.   Smith, E. & Morowitz, H. J. Universality in intermediary metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13168–13173 (2004).

13.   Cakmak, F. P. & Keating, C. D. Combining Catalytic Microparticles with Droplets Formed by Phase Coexistence: Adsorption and Activity of Natural Clays at the Aqueous/Aqueous Interface. *Sci. Rep.* **7**, 1–14 (2017).

14.   Mann, S. Systems of Creation: The Emergence of Life from Nonliving Matter. *Acc. Chem. Res.* **45**, 2131–2141 (2012).

15.   Dzieciol, A. J. & Mann, S. Designs for life: Protocell models in the laboratory. *Chem. Soc. Rev.* **41**, 79–85 (2012).

16.   Ruiz-Mirazo, K., Briones, C. & Escosura, A. de la. Prebiotic Systems Chemistry: New Perspectives for the Orgins of Life. *Chem. Rev.* **114**, 285–366 (2014).

17.   Rao, M., Eichberg, J. & Oro, J. Synthesis of Phosphatidylcholine under Possible Primitive Earth Conditions. *J Mol Evol* **18**, 196–202 (1982).

18.   Hargreaves, W. R., Mulvihill, S. J. & Deamer, D. W. Synthesis of phospholipids and membranes in prebiotic conditions. *Nature* **266**, 78–80 (1977).

19.   Fiore, M. & Strazewski, P. Prebiotic Lipidic Amphiphiles and Condensing Agents on the Early Earth. *Life* **6**, 17 (2016).

20.   Milshteyn, D., Damer, B., Havig, J. & Deamer, D. Amphiphilic compounds assemble into membranous vesicles in hydrothermal hot spring water but not in seawater. *Life* **8**, (2018).

21.   Mansy, S. S. Model protocells from single-chain lipids. *Int. J. Mol. Sci.* **10**, 835–

843 (2009).

22. Monnard, P.-A. & Walde, P. Current ideas about prebiological compartmentalization. *Life* **5**, 1239–1263 (2015).

23. Rushdi, A. I. & Simoneit, B. R. T. Lipid formation by aqueous Fischer-Tropsch-type synthesis over a temperature range of 100 to 400 °C. *Orig. Life Evol. Biosph.* **31**, 103–118 (2001).

24. Thomas, J. A. & Rana, F. R. The influence of environmental conditions, lipid composition, and phase behavior on the origin of cell membranes. *Orig. Life Evol. Biosph.* **37**, 267–285 (2007).

25. Sarkar, S., Dagar, S., Verma, A. & Rajamani, S. Compositional heterogeneity confers selective advantage to model protocellular membranes during the origins of cellular life. *Sci. Rep.* **10**, 4483 (2020).

26. Rajamani, S. *et al.* Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Orig. Life Evol. Biosph.* **38**, 57–74 (2008).

27. Olasagasti, F. & Rajamani, S. Lipid-assisted polymerization of nucleotides. *Life* **9**, 1–10 (2019).

28. Mungi, C. V., Rajamani, S., Mungi, C. V. & Rajamani, S. Characterization of RNA-Like Oligomers from Lipid-Assisted Nonenzymatic Synthesis: Implications for Origin of Informational Molecules on Early Earth. *Life (Basel, Switzerland)* **5**, 65–84 (2015).

29. Rodriguez-Garcia, M. *et al.* Formation of oligopeptides in high yield under simple programmable conditions. *Nat. Commun.* **6**, 1–7 (2015).

30. Varma, S. J., Muchowska, K. B., Chatelain, P. & Moran, J. Native iron reduces CO2 to intermediates and end-products of the acetyl-CoA pathway. *Nat. Ecol. Evol.* **2**, 1019–1024 (2018).

31. Muchowska, K. B. *et al.* Metals promote sequences of the reverse Krebs cycle. *Nat. Ecol. Evol.* **1**, 1716–1721 (2017).

32. Messner, C. B., Driscoll, P. C., Piedrafita, G., De Volder, M. F. L. & Ralser, M. Nonenzymatic gluconeogenesis-like formation of fructose 1,6-bisphosphate in ice. *Proc. Natl. Acad. Sci.* **114**, 7403–7407 (2017).

33. Coggins, A. J. & Powner, M. W. Prebiotic synthesis of phosphoenol pyruvate by

α-phosphorylation-controlled triose glycolysis. *Nat. Chem.* **2**, 1–76 (2016).

34. Namani, T. & Walde, P. From decanoate micelles to decanoic acid/dodecylbenzenesulfonate vesicles. *Langmuir* **21**, 6210–6219 (2005).

35. Duhan Toparlak, Ö. *et al.* Cyclophospholipids increase protocellular stability to metal ions. *J. Chem. Inf. Model.* **53**, 1–8 (2013).

36. Sarkar, S., Dagar, S., Lahiri, K. & Rajamani, S. pH-responsive self-assembled compartments as tuneable model protocellular membrane systems. (2022).

37. Jin, L., Kamat, N. P., Jena, S. & Szostak, J. W. Fatty Acid/Phospholipid Blended Membranes: A Potential Intermediate State in Protocellular Evolution. *Small* **1704077**, 1–9 (2018).

38. Szostak, J. W., Bartel, D. P. & Luisi, P. L. Synthesizing life. *Nature* **409**, 387–390 (2001).

39. Gilbert, W. Origin of life: The RNA world. *Nature* **319**, (1986).

40. Pasek, M. A. Schreibersite on the early earth: Scenarios for prebiotic phosphorylation. *Geosci. Front.* **8**, 329–335 (2016).

41. Pasek, M. A role for phosphorus redox in emerging and modern biochemistry. *Curr. Opin. Chem. Biol.* **49**, 53–58 (2019).

42. Dagar, S., Sarkar, S. S. & Rajamani, S. Geochemical influences on nonenzymatic oligomerization of prebiotically relevant cyclic nucleotides. *RNA* **26**, 756–769 (2020).

43. Ferris, J. P. Mineral Calalysis and Prebiotic Synthesis : Montmorillonite-Catalysed Formation of RNA. *Elements* **1**, 145–149 (2005).

44. Lailach, G., Thompson, T., Miner, G. B.-C. C. & 1968, U. Absorption of pyrimidines, purines, and nucleosides by Co-, Ni-, Cu-, and Fe (III)-montmorillonite (Clay-organic studies XIII). *Clays Clay Miner.* **16**, 295–301 (1968).

45. Joshi, P. C. & Aldersley, M. F. Significance of mineral salts in prebiotic RNA synthesis catalyzed by montmorillonite. *J. Mol. Evol.* **76**, 371–379 (2013).

46. Li, L. *et al.* Enhanced Nonenzymatic RNA Copying with 2-Aminoimidazole Activated Nucleotides. *J. Am. Chem. Soc.* **139**, 1810–1813 (2017).

47. Yi, R., Hongo, Y. & Fahrenbach, A. C. Synthesis of imidazole-activated ribonucleotides using cyanogen chloride. *Chem. Commun.* 1–4 (2018).

doi:10.1039/C7CC08489G

48. Bapat, N. V. & Rajamani, S. Templated replication (or lack thereof) under prebiotically pertinent conditions. *Sci. Rep.* **8**, 15032 (2018).

49. Kaddour, H. & Sahai, N. Synergism and mutualism in non-enzymatic RNA polymerization. *Life (Basel, Switzerland)* **4**, 598–620 (2014).

50. Rajamani, S. *et al.* Effect of Stalling after Mismatches on the Error Catastrophe in Nonenzymatic Nucleic Acid Replication. *J. Am. Chem. Soc.* **132**, 5880–5885 (2010).

51. Bapat, N. V. & Rajamani, S. Effect of Co-solutes on Template-Directed Nonenzymatic Replication of Nucleic Acids. *J. Mol. Evol.* **81**, 72–80 (2015).

52. Zhang, W., Pal, A., Ricardo, A. & Szostak, J. W. Template-Directed Nonenzymatic Primer Extension Using 2-Methylimidazole-Activated Morpholino Derivatives of Guanosine and Cytidine. *J. Am. Chem. Soc.* jacs.9b06453 (2019). doi:10.1021/jacs.9b06453

53. Mansy, S. S. *et al.* Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454**, 122–125 (2008).

54. Walton, T., Zhang, W., Li, L., Tam, C. P. & Szostak, J. W. The Mechanism of Nonenzymatic Template Copying with Imidazole-Activated Nucleotides. *Angew. Chemie Int. Ed.* **58**, 10812–10819 (2019).

55. Hagenbuch, P., Kervio, E., Hochgesand, A., Plutowski, U. & Richert, C. Chemical Primer Extension: Efficiently Determining Single Nucleotides in DNA. *Angew. Chemie Int. Ed.* **44**, 6588–6592 (2005).

56. Lawrence, M. S. & Bartel, D. P. Processivity of Ribozyme-Catalyzed RNA Polymerization. *Biochemistry* **42**, 8748–8755 (2003).

57. Walton, T., Dasgupta, S., Duzdevich, D., Soo, S. & Szostak, J. W. In vitro selection of ribozyme ligases that use prebiotically plausible 2-aminoimidazole – activated substrates. *PNAS* 1–8 (2020). doi:10.1073/pnas.1914367117

58. Horning, D. P. & Joyce, G. F. Amplification of RNA by an RNA polymerase ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **21**, 1–6 (2016).

59. Freisinger, E. & Sigel, R. K. O. From nucleotides to ribozymes-A comparison of their metal ion binding properties. *Coord. Chem. Rev.* **251**, 1834–1851 (2007).

60.	Young, K. J., Gill, F. & Grasby, J. A. Metal ions play a passive role in the hairpin ribozyme catalysed reaction. *Nucleic Acids Res.* **25**, 3760–3766 (1997).

61.	Sigel, R. K. O. Group II intron ribozymes and metal ions - A delicate relationship. *Eur. J. Inorg. Chem.* 2281–2292 (2005). doi:10.1002/ejic.200401007

62.	Trawick, B. N., Daniher, A. T. & Bashkin, J. K. Inorganic Mimics of Ribonucleases and Ribozymes: From Random Cleavage to Sequence-Specific Chemistry to Catalytic Antisense Drugs. *Chem. Rev.* **98**, 939–960 (1998).

63.	Brion, P. & Westhof, E. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 113–137 (1997).

64.	Athavale, S. S. *et al.* RNA Folding and Catalysis Mediated by Iron (II). *PLoS One* **7**, e38024 (2012).

65.	Wächtershäuser, G. On the Chemistry and Evolution of the Pioneer Organism. *Chem. Biodivers.* **4**, 584–602 (2007).

66.	Martin, W. F., Sousa, F. L. & Lane, N. Energy at life's origin. *Science (80-.).* **344**, 1092–1093 (2014).

67.	Liu, J. *et al.* Metalloproteins containing cytochrome, iron-sulfur, or copper redox centers. *Chem. Rev.* **114**, 4366–4369 (2014).

68.	Raanan, H., Poudel, S., Pike, D. H., Nanda, V. & Falkowski, P. G. Small protein folds at the root of an ancient metabolic network. *Proc. Natl. Acad. Sci.* **117**, 7193–7199 (2020).

69.	Keller, M. A., Turchyn, A. V. & Ralser, M. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. *Mol. Syst. Biol.* **10**, 1–12 (2014).

70.	Keller, M. A. *et al.* Conditional iron and pH-dependent activity of a non-enzymatic glycolysis and pentose phosphate pathway. *Sci. Adv.* **2**, e1501235 (2016).

71.	Belmonte, L. & Mansy, S. S. Metal catalysts and the origin of life. *Elements* **12**, 413–418 (2016).

72.	Wächtershäuser, G. Iron-Sulfur World. in *Wiley Encyclopedia of Chemical Biology* 1–8 (John Wiley & Sons, Inc., 2008). doi:10.1002/9780470048672.wecb264

73.	Wächtershäuser, G. Groundworks for an evolutionary biochemistry: The iron-sulphur world. *Prog. Biophys. Mol. Biol.* **58**, 85–201 (1992).

74. Wang, W. & Yang, B. FeS/S/FeS 2 Redox System and Its Oxidoreductase-like Chemistry in the Iron-Sulfur World 1 1. **11**, (2011).

75. Mateo-Marti, E., Galvez-Martinez, S., Gil-Lozano, C. & Zorzano, M.-P. Pyrite-induced uv-photocatalytic abiotic nitrogen fixation: implications for early atmospheres and Life. *Sci. Rep.* **9**, 15311 (2019).

76. Dalai, P. & Sahai, N. A Model Protometabolic Pathway Across Protocell Membranes Assisted by Photocatalytic Minerals. *J. Phys. Chem. C* **124**, 1469–1477 (2020).

77. Goldford, J. E., Hartman, H., Smith, T. F. & Segr??, D. Remnants of an Ancient Metabolism without Phosphate. *Cell* **168**, 1126-1134.e9 (2017).

78. Calvin, M. Evolution of Enzymes and the Photosynthetic Apparatus. *Science (80-.).* **130**, 1170–1174 (1959).

79. Kim, J. D. *et al.* Minimal Heterochiral de Novo Designed 4Fe–4S Binding Peptide Capable of Robust Electron Transfer. *J. Am. Chem. Soc.* **140**, 11210–11213 (2018).

80. Bonfio, C. *et al.* UV-light-driven prebiotic synthesis of iron-sulfur clusters. *Nat. Chem.* **9**, 5907 (2017).

81. Scintilla, S. *et al.* Duplications of an iron-sulphur tripeptide leads to the formation of a protoferredoxin. *Chem. Commun.* **52**, 13456–13459 (2016).

82. Bonfio, C. & Mansy, S. S. The Chemical Roots of Iron–Sulfur Dependent Metabolism. *Biochemistry* **56**, 5225–5226 (2017).

83. Bonfio, C. *et al.* Prebiotic iron–sulfur peptide catalysts generate a pH gradient across model membranes of late protocells. *Nat. Catal.* **1**, 616–623 (2018).

84. Raanan, H., Pike, D. H., Moore, E. K., Falkowski, P. G. & Nanda, V. Modular origins of biological electron transfer chains. *Proc. Natl. Acad. Sci.* **115**, 1280–1285 (2018).

85. Pandey, S. *et al.* Ladakh: Diverse, high-altitude extreme environments for off-earth analogue and astrobiology research. *Int. J. Astrobiol.* 1–21 (2019). doi:10.1017/S1473550419000119

Chapter 2

Geochemical influences on nonenzymatic oligomerization of prebiotically relevant cyclic nucleotides

(Adapted from, Dagar et.al. 2020; RNA)

2.1 Introduction

The "RNA World hypothesis" highlights the potential of RNA as the first biomolecule to have emerged on the prebiotic Earth. This is due to the capability of RNA to act as a catalyst in addition to being a genetic material [1]. The presence of complex protein machinery that could have assisted in the process of encoding and processing genetic information would have been highly improbable on early Earth. Therefore, the formation and replication of RNA molecules on prebiotic Earth would have been nonenzymatic; driven by the environmental conditions and the physio-chemical properties of the molecules, resulting in oligonucleotides. Small oligoribonucleotides (up to trimers) have been shown to not only enhance the rate of template-directed replication [2–5], but are also thought to act as precursors for cofactors that pervade extant biology [6–8].

Towards this, the prebiotically plausible synthesis of the components of RNA i.e., nucleotides, has been demonstrated. However, initial studies pertaining to the formation of RNA polymers and their propagation, have predominantly employed chemically activated nucleotides [9,10]. These activated nucleotides consist of a good leaving group, the presence of the leaving group enhances the electrophilicity of the phosphate group and hence facilitates the nucleophilic attack by the 2´ or 3´ hydroxyl group of an incoming nucleotide (Figure 1.1.). Though recently, a prebiotically plausible synthesis pathway for such activated nucleotides has been demonstrated [11]. Owing to their susceptibility to high temperature and aqueous environment, their presence in significant amounts on prebiotic Earth is uncertain, thus their potential to act as a substrate for enzyme-free oligomerization is questionable.

Towards this, the lipid-assisted non-activated nucleotides such as nucleoside 5´-monophosphate (NMPs) were shown to oligomerize in RNA-like oligomers [12]. These reactions were facilitated under acidic terrestrial geothermal conditions i.e., highly acidic pH (pH~2) and high temperature. However, systematic characterization of these polymers has shown oligomers with abasic sites due to depurination under such

conditions [13]. Hence, it is disputable that the resultant moieties that formed under such conditions would have been capable of functioning as informational oligomers.

As mentioned in Chapter 1, all the aforementioned reasons necessitate the requirement of intrinsically reactive, prebiotically relevant nucleotides. One such candidate is cyclic nucleoside monophosphates (cNMPs). These nucleotides are shown to be formed readily by dry heating of NMPs and also have been detected in mineral-mediated phosphorylation of nucleosides [14]. Their plausibility to form readily under prebiotic conditions and stability to high temperatures while being intrinsically active [14–20], highlights their potential to serve as monomers for enzyme-free oligomerization reactions, hence, resulting in RNA. Depending on the groups involved in the intramolecular phosphodiester bond ring, two types of cNMP are known i.e., 2´-3´cNMP and 3´-5´cNMP (2´-3´ and 3´-5´ groups are involved in the ring, respectively). Earlier studies have demonstrated the nonenzymatic oligomerization of 2′–3′ cNMPs under completely dry conditions [15]. Subsequent studies from di Mauro's group have shown the formation of RNA oligomers up to tetramers using 3′–5′ cNMP under dry and slightly alkaline regimes [21,22]. Dry conditions can provide a concentration mechanism, thus allowing for the reactants to interact with each other by bringing them into closer proximity. Thereby, such regimes facilitate the oligomerization, while decreasing the back-hydrolysis of oligomers formed, if any. Although insightful, an important aspect pertaining to these studies is the uncertainty of the existence of permanently dry niches on the early Earth. The realistic approach would then be to consider these reactions in niches that have been hypothesized to support the chemical emergence of life. One of these proposed niches is that of terrestrial hydrothermal pools. Temperature and seasonal fluctuations in these terrestrial pools would result in dry-wet cycles (Dehydration–Rehydration/DH–RH cycles) [23,24], a feature common to the geology of the planet. Such DH–RH cycles have been demonstrated to facilitate the formation of a diverse set of RNA building blocks (purine nucleosides) [25,26], and also in the condensation of hydroxy acids, either alone or in combination with amino acids, to result in oligoesters and depsipeptides, respectively [27,28]. As mentioned earlier, lipids have been shown to facilitate the oligomerization of 5′-NMPs under DH–RH conditions, resulting in the formation of RNA-like oligomers (abasic oligomers) [12,13]. Upon

dehydration lipid molecules arrange themselves into two-dimensional arrays, concentrating the reacting monomers thereby promoting concentration-dependent oligomerization. Upon rehydration, the multilamellar structures can encapsulate the monomers, as well as the growing oligomers, forming protocell-like entities, thus protecting them from hydrolysis [29,30]. Despite their implicit relevance, the Influence of lipids on enzyme-free oligomerization of cNMPs has not been explored till date. Additionally, most of the previous studies that have been undertaken to demonstrate the nonenzymatic oligomerization of cyclic nucleotides have used purines as the starting monomer [15,16,22]. Very few studies have looked at the oligomerization of pyrimidines under dry conditions [14,17]. Purines are known for their tendency to stack better than pyrimidines. This increases the proximity of these nucleotides thereby facilitating intermolecular condensation leading to their efficient oligomerization [31]. Given this, we aimed to elucidate whether there is any effect from the aforesaid difference (between purine and pyrimidine) on the oligomerization efficiency under our reaction conditions. Furthermore, we also wanted to discern the role of lipid molecules on the oligomerization of cyclic nucleotides under prebiotically pertinent DH–RH regimes. Toward this, enzyme-free oligomerization of cyclic nucleotides was studied using cyclic adenosine monophosphate (cAMP), a purine, and cyclic cytidine monophosphate (cCMP), a pyrimidine nucleotide, using both the 2′–3′ and 3′–5′ cyclic isomer versions. Lastly, the studies pertaining to nucleotide oligomerization are predominantly carried out under laboratory-controlled conditions. Although insightful, such controlled reactions may not be representative of the complex scenarios that might have been present on the early Earth. Toward this, few studies have begun to evaluate prebiotic processes under more "realistic" conditions. For example, Deamer et al. attempted to constrain the range of possible environments/conditions for the origins of life by analyzing prebiotically pertinent processes under realistic conditions [32]. The previous study from our laboratory looked at the self-assembly of simple amphiphiles, like fatty acids and their derivatives, in alkaline hot spring water samples of varying ionic content that were collected from three locations in Ladakh, a site with early Earth-like environmental niches [33]. However, to our knowledge, similar studies have not been undertaken in the context of nonenzymatic oligomerization of nucleotides, under prebiotic early Earth

environments. Therefore, we also set out to investigate how cyclic nucleotide oligomerization reactions would advent in a complex

"realistic" context. We evaluated this by using a hot spring sample that was used in the previous study [33], and studied it alongside reactions carried out under laboratory-controlled conditions. Under the laboratory-simulated conditions, intact oligomers up to trimers (in the pyrimidine reactions) and tetramers (in the purine reactions) were observed in some of the oligomerization reactions. Interestingly, in the reactions performed using the hot spring water sample, not only the rate of oligomerization but also the destabilization of the oligomers was enhanced over prolonged cycles.

2.2 Materials and Methods

2.2.1 Materials

The monosodium salts of all four cyclic monophosphates viz. adenosine 3′,5′-cyclic monophosphate (3′,5′-cAMP), adenosine 2′,3′-cyclic monophosphate (2′,3′-cAMP), cytidine 3′,5′-cyclic monophosphate (3′,5′-cCMP), cytidine 2′,3′-cyclic monophosphate (2′,3′-cCMP), adenosine monophosphate (AMP) linear dimer, trimer, and tetramer were purchased from Sigma–Aldrich and used without further purification. 1-palmitoyl-2-oleoyl-sn- glycero-3-phosphocholine (POPC) was purchased from Avanti Polar Lipids Inc. All other reagents used were of analytical grade and purchased from Sigma–Aldrich. TLC Silica gel 60 F254 was purchased from Merck (EMD Millipore Corporation).

2.2.2 Methods

2.2.2.1. Simulating early Earth conditions

Early Earth conditions were simulated using a bench-top heating block that was maintained at high temperatures, that is, 90°C as described by Mungi et al.[13]. The oligomerization reactions were carried out in 20 ml glass vials with their caps fitted with PTFE septa purchased from Chemglass. The anoxygenic environment was maintained by delivering a gentle flow of carbon dioxide into the vials through two

PEEK tubings of about 1–1.5 in, one acting as an inlet and another as an outlet for the gas.

2.2.2.2. Vesicle solution preparation

The vesicle solutions were prepared by drying the required amount of chloroform solution (with strength 25 mg/ml) of POPC under nitrogen gas flow to prepare a dry lipid film. It was then kept under vacuum for 5–6 h to make sure that no trace amount of chloroform remained. Subsequently, nanopure water was used to rehydrate the thin film to form the vesicles.

2.2.2.3. Nonenzymatic oligomerization reaction conditions and procedure

A typical reaction mixture consisted of 5 mM cyclic nucleotide in nanopure water with pH ~8.5 or the hot spring water sample. In lipid-assisted reactions, 1 mM of POPC was added to each reaction to maintain a constant ratio of 5:1 for nucleotide to lipid. The lipid concentration was chosen based on a previous study in our laboratory [13]. The samples were rehydrated with nanopure water and mixed properly followed by a prolonged dehydration phase during each DH–RH cycle, with 24 h per cycle. The rehydrated samples were collected at different periods. The lipid was extracted from the rehydrated sample using a previously standardized butanol-hexane extraction procedure. The aqueous portion containing nucleotides was collected and analyzed using HPLC and LC-MS/MS.

2.2.2.4. Microscopic analysis

Ten microliters of the reaction sample containing only 1 mM of POPC, was placed on a glass slide which was then evenly spread and covered with a 18 X 18mm coverslip. This was then air sealed by covering its four sides with liquid paraffin. This was done in order to decrease the motion of solution. The slides prepared were then observed under 40X magnification using a DIC microscope AxioImager Z1, (NA= 0.75) to check for the presence of lipid vesicles.

2.2.2.5. HPLC analysis

HPLC analysis was performed using the Infinity series 1260 HPLC instrument (Agilent Technologies, Santa Clara, CA, USA). The reaction products, after

removing the lipid molecules using previously standardized butanol-hexane extraction, were loaded onto HPLC [13]. The molecules were separated by an anion-exchange column viz. DNAPac PA 200 (Thermo Scientific, Sunnyvale, CA, USA). It separates the molecules based on their interaction with the column through phosphate moiety, thus providing single-nucleotide resolution. The separation method was standardized using a gradient of 2 mM Tris buffer at pH 8 (Solvent A) and 0.4 M sodium perchlorate in 2 mM Tris buffer at pH 8 as an eluting solvent (Solvent B). The gradient used was: 0% solvent B for 3 minutes, from 3 min to 10 min, solvent B goes from 0 to 30%, in next 3 min (10-13 min) it increases to 100%, where it is kept for 3 min (13-16 min), followed by equilibration with solvent A from 18-21 min. A photo Diode-Array Detector (DAD) was used to detect analytes at 260nm, using a highly sensitive 60 mm flow cell. The separation between a dead volume peak (which was later characterized as open ring structures lacking phosphate moiety, i.e., adenosine or adenine by mass), cAMP monomer, linear AMP monomer, and oligomers was observed in a typical HPLC chromatogram. The analytes were semi-quantified by measuring the area under corresponding peaks separated through column. Although the semi-quantified area can be due to the presence of various species with the same charge, it can still provide us with the distribution of species in the reaction after a particular DH-RH cycle.

2.2.2.6. Mass analysis

Mass analysis of the samples was performed on a Sciex X500R QTOF mass spectrometer (MS) fitted with an Exion-LC series UHPLC (Sciex) using information-dependent acquisition (IDA) scanning method. The acquired data were analyzed using the Sciex OS software (Sciex; University of Florida). The crude reaction mixture was separated on Zorbax C8 column (dimensions: 4.6×150 mm, 3.6 µm particle size) (Thermo Scientific) fitted with its guard column. A gradient of nanopure water containing 0.1% formic acid and acetonitrile containing 0.1% formic acid was used to separate the oligomers. The gradient used was: 0% solvent B for 3 minutes, from 3 min to 7 min, solvent B goes from 0 to 20%, in next 9 min (7-16 min) it increases to 30%, followed by a further increase to 100% from 16-18 min, the column is kept at 100% solvent B for next 3 min (18-21 min), followed by equilibration with solvent A

from 22-25 min. All the mass acquisition was performed using electron spray ionization (ESI) with the following parameters: turbo spray ion source, medium collision gas, curtain gas=30 L/min, ion spray voltage=5500 V (positive mode), at 500°C. TOF-MS acquisition was done at declustering potential as 80 V with 20 V as spread and using 10 V collision energy. To perform TOF-MS–MS analysis, 50 V collision energy with 20 V spread was used. As the mass acquisition was carried out in positive mode, the observed masses correspond to the mass of the H+-adduct of the parent molecule. The presence of a specific species/molecule was confirmed by the presence of precursor mass within the 5-ppm error range as well as its fragmentation pattern.

2.2.2.7. TLC (thin layer chromatography) of POPC

TLC analysis for evaluating the chemical stability of POPC molecules was performed using TLC Silica gel 60 F254 (Merck, EMD Millipore Corporation). A mixture of chloroform, methanol, and water in 65:25:4 was used as the mobile phase. The TLC chamber (Latch-lid TLC developing chambers for 10×10 cm plates, Sigma–Aldrich) was lined with filter paper and equilibrated with the solvent system before use. The TLC plates (dimension 5× 10 cm) were pre-run with the mobile phase in order to eliminate interference coming from any intrinsic contaminant. It was then dried, and 6 μL of each sample was withdrawn at different DH–RH cycles (cycle 0, cycle 4, cycle 10, cycle 20, and cycle 30 in lanes b, c, d, e, and f, respectively) along with POPC as a control (lane a), was spotted 0.8 cm apart to each other and 2 cm above the bottom edge of TLC plate. The dried TLC plates were then chromato- graphed in the equilibrated chamber. Following the chromatography, the TLC plates were dried and kept in the presaturated tank containing iodine vapours. The plates were then immediately scanned on a Syngene G-Box Chemi-XRQ gel documentation system (Syngene).

2.3. Results

2.3.1. Nonenzymatic oligomerization of cyclic adenosine monophosphate under DH–RH conditions

We investigated the enzyme-free oligomerization of cyclic nucleotides under laboratory-simulated DH–RH regimes. A typical reaction involved 5 mM of the corresponding cyclic nucleotides (2′–3′ or 3′–5′ cAMP) at pH 8, which were subjected to 30 DH– RH cycles (at 24 h per cycle) (Methods). A dry heating control reaction was also set up under identical temperature, pH and reactant concentrations, wherein the sample was only subjected to prolonged drying without any rehydration involved. This was also done to recreate conditions used in previous studies [15,16,21,22]. The reactions were analyzed using high-performance liquid chromatography (HPLC) (Methods). The product identities were ascertained by coeluting them on HPLC with known controls. The control species including cAMP, AMP monomer, dimer, and trimer, showed peaks eluting at 1.8, 6.1, 6.3, and 7.1 min, respectively. The peak observed with a retention time of 1.5 min is adenine, which is mentioned as a breakdown peak in the chromatogram (Figure 2.1A). Free adenine does not interact with an anion-exchange column like the one that was used for this analysis and elutes out of the column in the void volume.

As shown in the HPLC chromatogram (Figure 2.1B), the reaction mixture kept at 90°C for 30 d under complete dry heating conditions led to the hydrolysis of cAMP to its linear form (i.e., AMP), however, no oligomers were observed for both of the cyclic isomers (2′, 3′ cAMP and 3′, 5′ cAMP) (Figure 2.1B, D). Conversely, under DH–RH conditions, a peak, which had an identical retention time as that of the control dimer, was observed in the 2′, 3′ cAMP oligomerization reaction (Figure 2.1A). Nonetheless, there was no clear separation observed among the other peaks with a retention time more than that of AMP in the HPLC chromatogram. This could be, both, due to the relatively low yields of the reaction, and due to the presence of different types of product species. For example, the resultant products could be either cyclic or linear versions, and could also have varied internucleotide linkages, especially in the 2′, 3′ cNMP reactions (Figure 2.2). However, in the case of 3′, 5′ cAMP reactions under DH–RH conditions, oligomeric species was not detected using HPLC (Figure 2.1C). All the reactions were performed at least in replicate to validate the results.
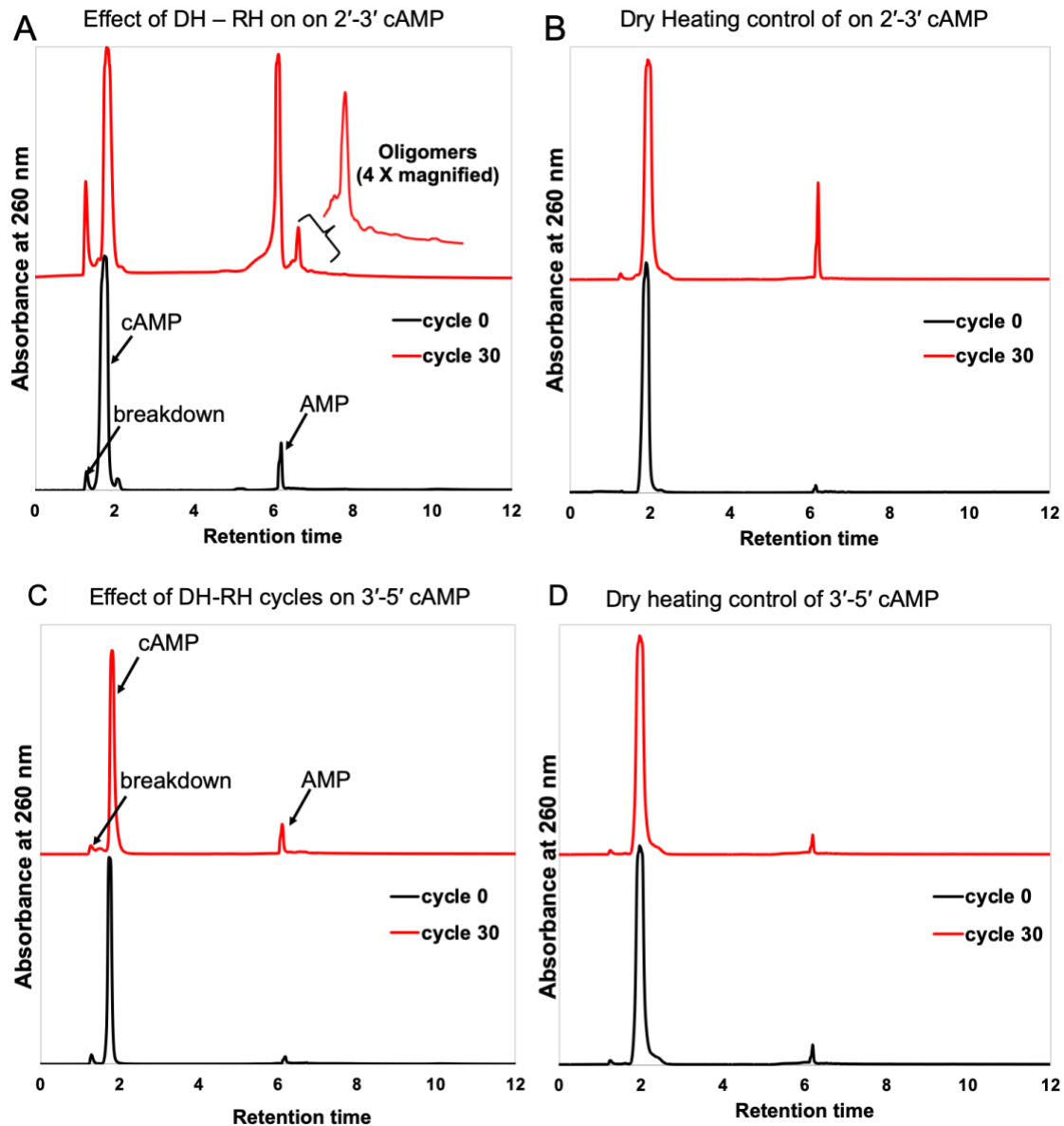
Figure 2.1. Effect of DH–RH cycles on the oligomerization of cAMP. Representative HPLC chromatograms for oligomerization reactions of 2′, 3′ cAMP and 3′, 5′ cAMP subjected to 30 DH–RH cycles (A and C, respectively); oligomerization of 2′, 3′ cAMP and 3′, 5′ cAMP, but only under prolonged dry heating (B and D, respectively). The black trace depicts the chromatogram for cycle 0, which was collected at time zero, and the red trace shows the chromatogram for reaction time point equivalent to 30 d. Y-axis shows the absorbance at 260 nm while the x-axis shows time in minutes.
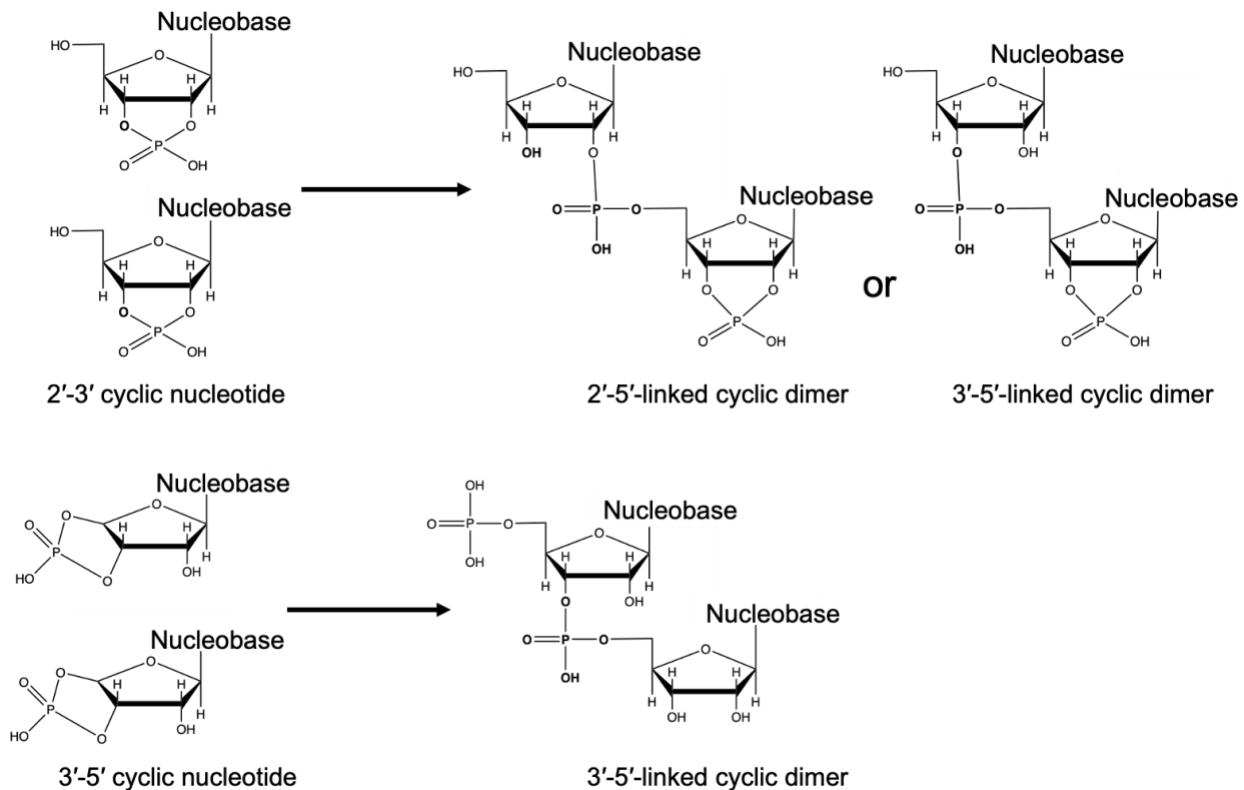
Figure 2.2. Schematic showing possible phosphodiester linked products resulting in the oligomerization reactions of 2′, 3′ cyclic nucleotides and 3′, 5′ cyclic nucleotides. Schematic made using ChemDraw Professional, version 19.0.0.26
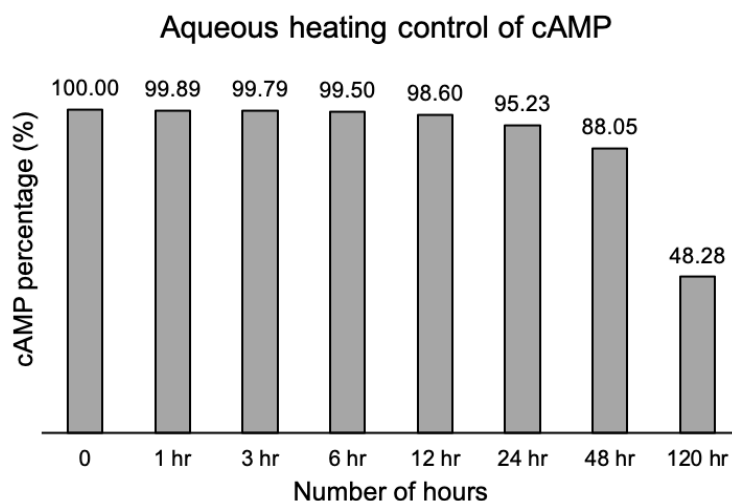
Figure 2.3. The stability of cAMP under aqueous heating. The graph shows the stability of the starting monomer (after accounting for both 2′, 3′ cAMP and AMP as they can

35

readily interchange in aqueous conditions). The X-axis shows different time periods (in hours) after which samples were withdrawn and analyzed using HPLC. Y-axis shows cAMP (%) remaining, normalized to the amount present in cycle zero.

Additionally, an aqueous heating control was also set up under identical conditions without invoking a dry phase and was monitored for 120 h. The stability of 2′, 3′ cAMP was compromised under such conditions and only breakdown (adenine) and AMP peaks were observed (48.28% of cAMP was remaining after 120 h of aqueous heating; Figure 2.3). Time of flight-mass spectrometry (TOF-MS) analysis was performed in order to further confirm and characterize the various species that were formed under DH–RH conditions and observed in the HPLC chromatograms (Methods). Fragmentation was induced for the precursor molecule using the TOF-MS/MS mode.
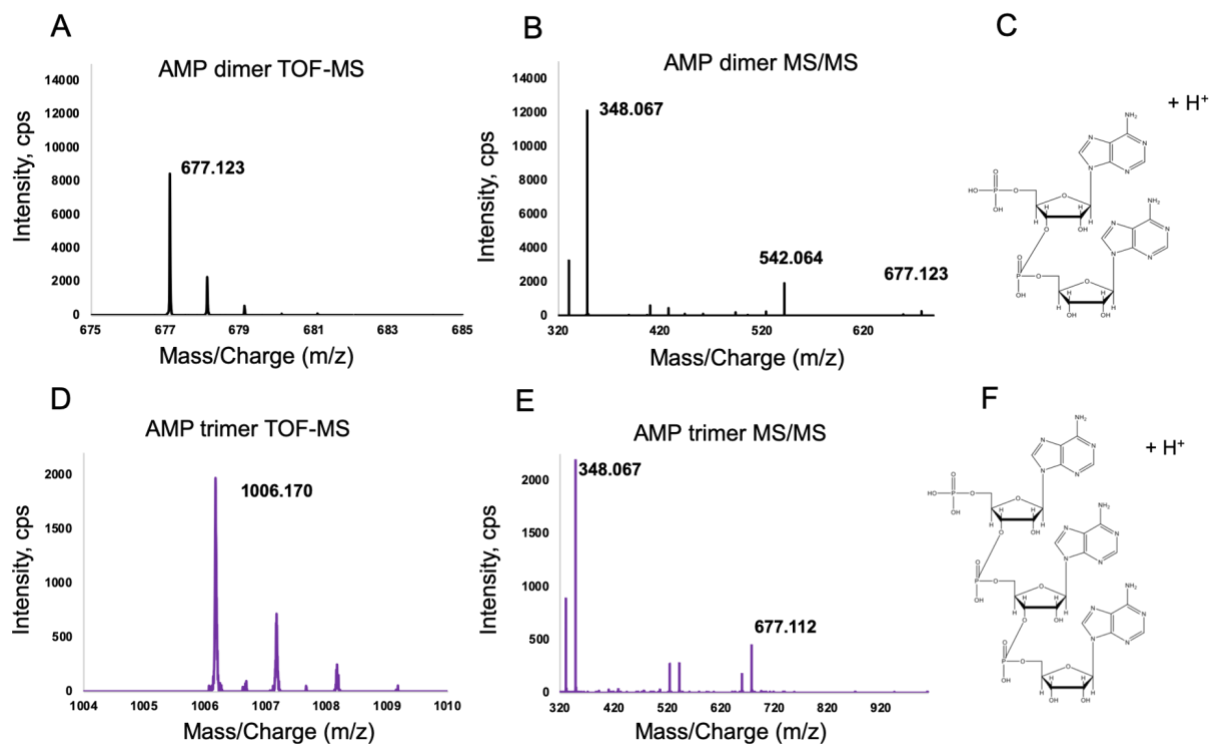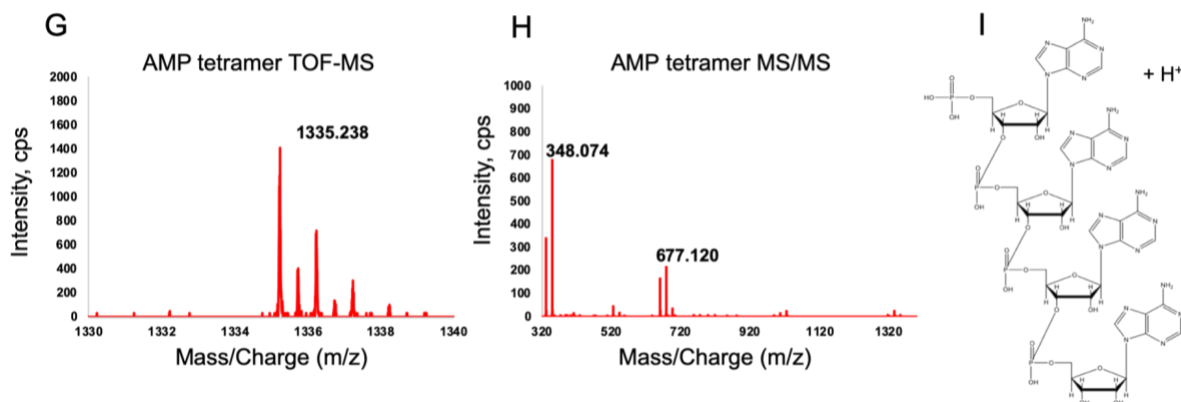
Figure 2.4. Representative TOF-MS and TOF-MS/MS spectra in ESI positive mode for cAMP reactions. From left to right: the TOF-MS spectrum, TOF-MS/MS spectrum, and the representative structure of different AMP oligomer species observed. The species correspond to H+ adducts of AMP linear dimer (A–C) with a monoisotopic mass of 677.12 Da, AMP linear trimer (D–F) with a monoisotopic mass of 1006.17 Da, and AMP linear tetramer (G–I) with a monoisotopic mass of 1335.23 Da, respectively.

Figure 2.4 shows representative TOF-MS and TOF-MS/MS spectra that were observed in electrospray ionization (ESI) positive mode for H+ adducts of AMP linear dimer (panels A and B), along with its corresponding structure (panel C), AMP linear trimer (panels D and E), and its corresponding structure (panel F), and AMP linear tetramer (panels G and H), and its corresponding structure (panel I). All the observed masses have been summarized in Table 2.1, with the errors being well within the acceptable range (give or take [±] 5 parts per million [ppm]). The fragmentation of the AMP linear dimer with a monoisotopic mass of 677.123 Da yielded an intense fragment at 348.067 Da, and another fragment at 542.064 Da (Figure 2.4A, B). AMP linear trimer with a monoisotopic mass of 1006.170 fragmented into two intense fragments of 348.067 Da, and 677.112 Da (Figure 2.4D, E). AMP linear tetramer with a monoisotopic mass of 1335.238 fragmented into two intense fragments of 348.074 Da, and 677.12 Da (Figure 2.4G, H). The fragmentation fingerprint of the parent molecule, its ppm accuracy, and the elution time were all used in conjunction to confirm the presence of the target molecules. The TOF MS spectrum for the 2′, 3′ cAMP reaction showed the presence of intact linear dimer within four DH–RH cycles,

and intact linear trimer and tetramer after 10 DH–RH cycles (Figures 2.4 and 2.8A; Table 2.4). However, 3′, 5′ cAMP yielded only an intact linear dimer even after 20 DH–RH cycles. This lower oligomerization propensity of 3′, 5′ cAMP (oligomers only up to dimer as compared to tetramer in the case of 2′, 3′ cAMP) could possibly stem from its higher intrinsic stability. More than 90% of the starting monomer typically remained even after 30 DH–RH cycles and this could possibly be due to comparatively less ring strain in its six-membered ring (8.9 kcal/mol) than in the five-membered ring (12 kcal/mol) of 2′, 3′ cAMP [34].

2.3.2. Nonenzymatic oligomerization of cyclic cytidine monophosphate under DH–RH conditions

In order to compare the oligomerization potential of purine versus pyrimidine cyclic nucleotides, reactions were carried out using cCMP. The reactant concentrations and reaction conditions including DH–RH cycling at 90°C (30 DH–RH cycles with 24 h per cycle), were kept identical to what was used in the cAMP reactions. Both of the cyclic isomers, that is, 2′, 3′-cCMP and 3′, 5′-cCMP were investigated for their nonenzymatic oligomerization potential. Mass analysis of the cCMP oligomerization reaction products was performed and the parameters for TOF-MS and TOF-MS/ MS were kept identical to that of the cAMP reactions (Methods). Figure 2.5 shows a representative TOF-MS spectra and TOF-MS/MS spectra in ESI positive mode that were observed for the H+ adduct of CMP linear dimer (panels A and B), along with its corresponding structure (panel C), and H+ adduct of CMP linear trimer (panels D and E) and its corresponding structure (panel F). As seen in Figure 2.5, the CMP linear dimer with a monoisotopic mass of 629.099 Da fragmented into daughter fragments with masses 611.087, 306.046, and 518.063 Da. CMP linear trimer with a monoisotopic mass of 934.142 fragmented into 306.047 Da, and 629.096 Da, respectively. All the observed masses with their respective ppm errors are summarized in Table 2.1. The TOF-MS and MS/MS analysis for the 2′, 3′ cCMP and 3′, 5′ cCMP reactions show the presence of linear trimer and linear dimer, respectively, within 10 DH–RH cycles (Figures 2.5 and 2.8B). The lower extent of

oligomerization for cCMP than cAMP (oligomers only up to trimer as compared to tetramer), could potentially be due to the intrinsically higher stability of cCMP toward the hydrolysis of intramolecular phosphodiester linkage ($\Delta H_{hydrolysis}$=−8.1 kcal/mol for cCMP and $\Delta H_{hydrolysis}$ =−9.4 kcal/mol for cAMP, both at pH 7.3, 25°C) [35].

Table 2.1. Summary of the masses in the nonenzymatic oligomerizations reactions of cAMP and cCMP.

	Species	Expected mass	Observed mass	Error (ppm)	Species	Expected mass	Observed mass	Error (ppm)
2′–3′ cNMP	cAMP	330.05	330.05	−0.90	cCMP	306.04	306.04	−1.30
	AMP dimer	677.12	677.12	2.51	CMP dimer	629.10	629.10	−0.15
	AMP trimer	1006.17	1006.17	1.78	CMP trimer	934.14	934.14	1.92
	AMP tetramer	1335.22	1335.23	2.62	CMP tetramer[a]	1239.18	1239.18	4.59
3′–5′ cNMP	cAMP	330.05	330.05	−1.51	cCMP	306.04	306.04	−0.98
	AMP dimer	677.12	677.12	1.77	CMP dimer	629.10	629.10	1.74

[a]Depicts that the species was observed in TOF-MS but not seen in TOF-MS–MS analysis.
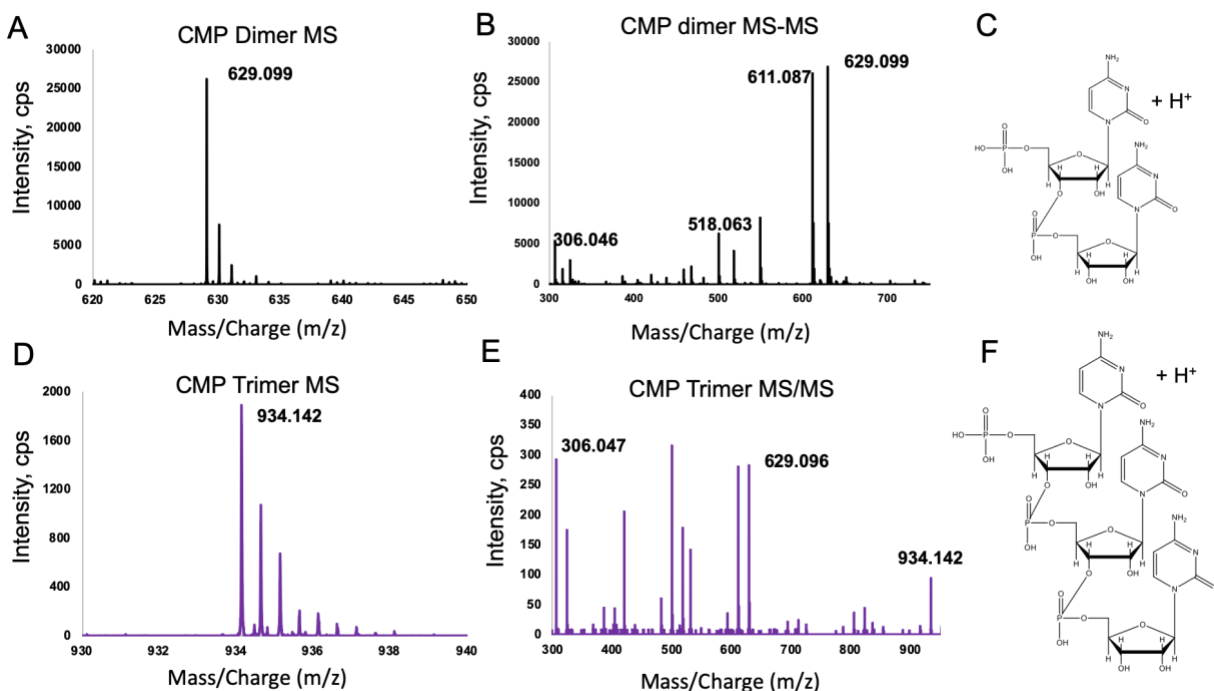
Figure 2.5. Representative spectra of TOF-MS and TOF-MS/MS in ESI positive mode for cCMP reactions. The figure represents the TOF-MS spectrum, TOF-MS/MS spectrum, and the structures that were observed for the H+ adducts of CMP linear

dimer (A–C) with monoisotopic mass 629.1 Da, and CMP linear trimer (D–F) with monoisotopic mass 934.14 Da, respectively.

2.3.3. Effect of lipids on the stability of cyclic nucleotides under DH–RH conditions

Lipids have been demonstrated to encapsulate macromolecules under DH–RH cycles [36]. They are also known to interact better with certain molecules present in the bulk solution [37,38], consequently protecting them from the harsh environmental conditions of an early Earth. In the context of nucleotide oligomerization reactions, lipids have been shown to also facilitate the formation of RNA-like polymers (abasic oligomers) from nucleoside-5′ monophosphates [12,13]. Given this, we set out to elucidate the plausible effect of lipids on the enzyme-free oligomerization of cNMP (cAMP and cCMP based reactions). The reactions were performed in the presence of 1-palmitoyl, 2-oleoyl phosphatidylcholine (POPC), a phospholipid that has been used as a proxy membrane component in several previous studies [12,13,39–41]. A total of 1 mM of POPC was added to each reaction containing 5mMof cNMP to maintain a constant ratio of 5:1 of nucleotide to lipid. The reaction conditions and subsequent procedures were kept identical to those mentioned in the previous section (90°C, 30 DH–RH cycles with 24 h per cycle) (Methods). In order to get an estimate of the amount of starting monomer in the HPLC runs, the areas under both linear NMP and cNMP peaks were taken into consideration, as cNMP can readily convert to NMP under aqueous conditions [20].

Figure 2.6 shows the graph of the semi-quantitation done on the HPLC data, which corresponds to the stability of the starting monomer (accounting for both cNMP and NMP). The amount of the starting monomer (either cAMP or cCMP) decreased with increasing DH–RH cycles (Figure 2.6; Table 2.2). In the 2′, 3′ cAMP reactions (without lipid) (Figure 2.6A), cAMP reduces to 80%, 50%, 30%, and 20% in 4, 10, 20, and 30 cycles, respectively. However, in lipid-assisted reactions, the starting monomer was more stable; up to 60% of the starting monomer persisted even after 30 DH– RH cycles. In the case of 3′, 5′ cAMP, in both lipid-assisted and lipid-free

reactions, only 5% of the starting monomer underwent breakdown even after 30DH–RH cycles. The results suggested that the presence of lipid enhanced the stability of the starting monomer in the 2′, 3′ cAMP reaction. However, the protecting effect of lipid was not evident in the reactions with 3′, 5′ cAMP as it is inherently much more stable.
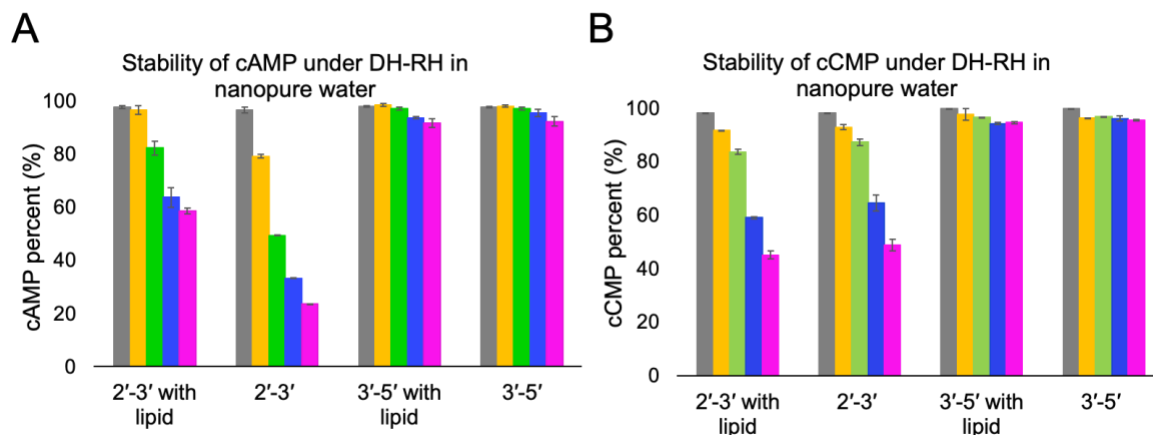
Figure 2.6. The stability of cyclic nucleotides in enzyme-free oligomerization reactions, in both, without lipid and lipid-assisted reactions. The graphs show the stability of the cAMP (A) and cCMP (B), during the reactions carried out under the DH–RH regimen, either in the presence or absence of lipids. The X-axis shows different reactions with indicated cNMP. Different colours depict different DH–RH cycles. Y-axis shows corresponding cNMP (%) remaining, normalized to that present in cycle zero. Error bars indicate SD.

In the case of 2′, 3′ cCMP reactions (Figure 2.6B; Table 2.2), the quantity of starting monomer reduced by 10%, 20%, 40%, and 55% in 4, 10, 20, and 30 d of DH–RH cycling, respectively, in both the without lipid and lipid-assisted reactions. 3′, 5′ cCMP was found to be more stable than the 2′, 3′ cCMP under our reaction conditions, similar to what was observed with cAMP isomers. In all, the concentration of 3′, 5′ cCMP reduced to 95% of its initial concentration in 30DH–RH cycles. Hence the lipid did not seem to have any influence on the cCMP reactions, both, in the case of 2′–3′ cCMP and 3′–5′ cCMP reactions. As alluded to earlier, the protecting effect of lipid on

a particular solute molecule could stem from either its entrapment inside a vesicle/within bilayers or by directly interacting with the amphiphilic surfaces. However, the mechanisms of interaction can vary for different molecules and might depend on the variables such as planarity, hydrophobicity, and hydrogen bonding [37]. Purines are shown to bind fatty acid membranes more than pyrimidines, mainly via hydrophobic interactions [37]. The reduced binding of pyrimidine nucleobases/nucleotides with lipid could possibly explain the negligible effect of the lipid in the case of nonenzymatic oligomerization reactions involving cCMP.

Table 2.2. Stability of cyclic nucleotides in enzyme-free oligomerization reactions, in both, without lipid and lipid-assisted reactions.

	cAMP				
	Cycle 0	Cycle 4	Cycle 10	Cycle 20	Cycle 30
2′–3′ with lipid	97.76 ± 0.62	96.55 ± 1.59	82.29 ± 2.66	63.66 ± 3.71	58.60 ± 1.14
2′–3′	96.59 ± 1.12	79.24 ± 0.61	49.32 ± 0.14	33.24 ± 0.13	23.61 ± 0.10
3′–5′ with lipid	98 ± 0.22	98.44 ± 0.51	97.26 ± 0.52	93.71 ± 0.44	91.70 ± 1.64
3′–5′	97.66 ± 0.24	98.11 ± 0.47	97.16 ± 0.49	95.58 ± 1.34	92.31 ± 1.73
	cCMP				
	Cycle 0	Cycle 4	Cycle 10	Cycle 20	Cycle 30
2′–3′ with lipid	98.25 ± 0.00	91.57 ± 0.24	83.61 ± 1.00	58.95 ± 0.46	45.09 ± 1.58
2′–3′	98.29 ± 0.00	92.89 ± 0.94	87.31 ± 1.24	64.51 ± 3.00	48.70 ± 2.21
3′–5′ with lipid	99.88 ± 0.02	97.69 ± 2.18	96.50 ± 0.12	94.24 ± 0.36	94.57 ± 0.47
3′–5′	99.95 ± 0.01	96.24 ± 0.02	96.85 ± 0.06	96.13 ± 1.00	95.51 ± 0.21

The table depicts the quantity (area percent under the HPLC peaks) of cAMP and cCMP that persisted after the indicated number of DH–RH cycles, either in the presence or absence of lipids. Error bars indicate SD.

2.3.4. Nonenzymatic oligomerization reactions of cyclic nucleotides in hot spring water

To obtain a more realistic picture of the oligomerization phenomenon involving cyclic nucleotides, the enzyme-free oligomerization reactions were also performed using water samples collected from a hot spring. Few previous studies have demonstrated that the presence of certain metal ions can catalyze nonenzymatic oligomerization reactions of nucleotides [42–44]. Thus, metal ions present in a hot spring could also potentially affect nonenzymatic oligomerization. Nonetheless, high amounts of salt are also known to be detrimental to RNA [45]. Panamic hot spring water that was used in the study by Joshi et al.[33] was chosen to investigate the aforementioned

oligomerization reactions in detail. This site was studied as part of the Spaceward Bound expedition to Ladakh, an astrobiologically relevant site in the northern part of India [46]. It is considered an early Earth and a Martian equivalent site due to its various topological features. Panamic water sample has a total salt content of around 7.6–7.9 mM, with predominantly Na^+ as major cation (~92.5%) followed by Ca^{2+} (~5%), K^+ (~2%), Li^+ (~0.7%), and Mg^{2+} (~0.03%) ions. HCO_3^- was found to be the major anion (~82%) followed by SO_4^{2-} (~14.3%) and Cl^- (~3.8%) (Table 2.3). The reaction conditions such as temperature and other parameters like reactants concentrations were kept identical to what was described for the aforesaid reactions that were studied in nanopure water (i.e., 30 DH–RH cycles, each cycle consists of 24 h at 90°C). The native pH of the hot spring water sample was ~8–8.5, which was kept as is in the experiments undertaken. Figure 2.8 gives a comparative abundance of the different oligomeric species (normalized to cNMP at cycle 0) that were observed in cAMP and cCMP both in the presence and absence of lipid (POPC), in nanopure and hot spring water reactions (Figure 2.8; Table 2.4). All the oligomers were characterized using LC-MS and MS/MS analysis. Lipid-free reactions of 2′, 3′ cAMP yielded oligomers up to intact AMP linear tetramer after four DH–RH cycles. However, it seemed to degrade after 10 cycles of DH–RH. AMP linear trimer was also observed to disappear after 20 DH–RH cycles. On the contrary, in lipid-assisted reactions, 2′, 3′ cAMP oligomerized to yield linear trimer in 4 d of DH– RH cycling and persisted even after 30 DH–RH cycles, indicating the stabilizing effect of lipid. However, in the case of 3′, 5′ cAMP reaction, only a dimer was observed after 10 DH–RH cycles, both, in the presence and absence of lipid. Therefore, the presence of lipid did not seem to make any significant difference in this case. In 2′, 3′ cCMP lipid-assisted and the without lipid reactions, CMP linear trimer was observed after 10 DH–RH cycles, which was stable throughout the reaction, that is, till 30 cycles. The experiments with 3′, 5′ cCMP yielded linear dimer after 20 DH– RH cycles. In general, the metal ions and other co-solutes (like bicarbonates, etc.) present in the hot spring water sample seemed to increase the rate of oligomerization reaction, albeit destabilizing the resultant oligomers too.

Table 2.3. Geochemical analysis of Panamic hot spring water sample.

Hot spring water	Major cations					Major anions					
	Na⁺	K⁺	Ca²⁺	Mg²⁺	Li⁺	HCO₃⁻	Cl⁻	SO₄²⁻	TZ⁺	TZ⁻	TZ⁺/TZ⁻
	All values in mM								μE		
Panamic	6.76	0.14	0.36	0.002	0.05	5.68	0.26	0.99	7665	7918	0.97

The net inorganic charge balance (TZ⁺/TZ⁻) around 1 indicates the high quality of analysis (adapted from Joshi et al. 2017).

2.3.5. Stability of POPC vesicles under multiple DH–RH cycles in nanopure vs Panamic water

Our results demonstrated that the presence of POPC could indeed enhance the stability of the starting monomer and oligomers in some of the reactions that were investigated (2′, 3′ cAMP in both nanopure and Panamic water) under DH–RH conditions. This corroborates with previous studies that have reported a similar protective effect coming from the presence of lipids in the reaction milieu [12,13]. Lipid bilayers can do so by trapping the substrate molecules in the interlamellar spaces during the dry phase [30]. Coincidentally, this also provides a concentration effect thereby facilitating uphill forward reactions like oligomerization. Furthermore, the lipids would also encapsulate the resultant oligomers upon rehydration, thus protecting them from back reactions like hydrolysis, which leads to their accumulation over multiple cycles. However, to the best of our knowledge, none of the previous studies has examined the stability of the POPC vesicles under multiple DH–RH cycles for a prolonged duration, for example, for 30 d, as was done in this study. In order to evaluate whether the lipid vesicles are stable under our reaction conditions, 1 mM of POPC was used and DH–RH cycling was performed, both, in nanopure water and Panamic water samples at pH 8 for 30 d. The stability of the POPC vesicles was monitored by observing them under a microscope after a different number of DH–RH cycles using differential interference contrast (DIC) microscopy (Methods).
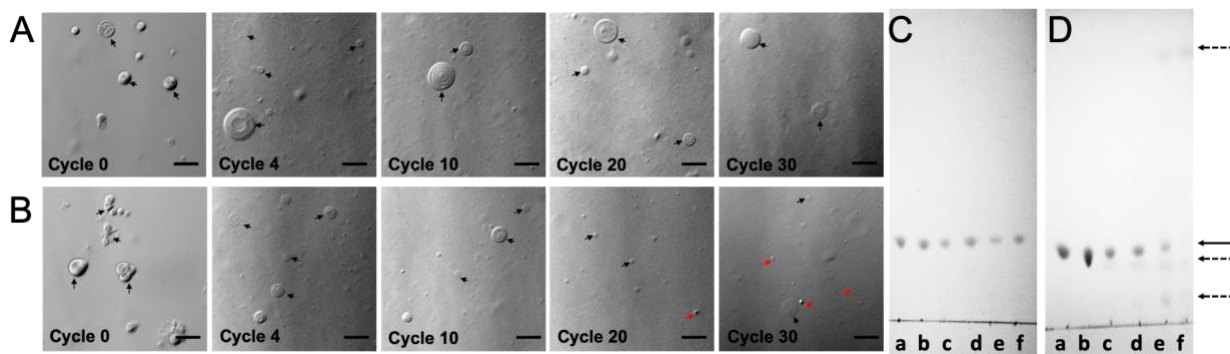
Figure 2.7. The stability of POPC vesicles under our reaction conditions. The stability of POPC vesicles under oligomerization reaction conditions with increasing DH–RH cycles (left to right), viz. cycle 0, cycle 4, cycle 10, cycle 20, cycle 30, either in nanopure water (A) or in Panamic water sample (B). The black and red arrows indicate vesicles and aggregates (collapsed vesicles), respectively. The scale bar in all the images is 10 μm. The chemical stability of POPC molecules under DH–RH conditions in nanopure water (C) and Panamic water (D). Different lanes from left to right, that is, a to f, indicate POPC control, the reactions corresponding to DH–RH cycle 0, cycle 4, cycle 10, cycle 20, and cycle 30, respectively. A solid black arrow indicates the band corresponding to POPC, dashed black arrows indicate new bands that appear corresponding to degradation products.

Vesicles were observed in nanopure water reactions even after 30 DH–RH cycles (Figure 2.7A). On the contrary, small aggregates (collapsed vesicles) were observed along with vesicles after 20 and 30 DH–RH cycles in the Panamic water sample (Figure 2.7B). Nonetheless, an important aspect to consider here is that the mere presence of vesicles does not necessarily have to be a direct reflection of the molecular stability of POPC. Previous literature in the field has demonstrated that the fatty acid chains can also self-assemble into vesicles around the pH that is near or equal to their pKa [47]. Oleic acid has a pKa around 8.5, consequently, at our reaction pH, that is, 8, if POPC is degrading to its components, the resulting oleic acid can potentially form vesicles. Thus, the vesicles observed could also be a false positive indication of the stability of POPC. Therefore, the chemical stability of POPC molecules was also evaluated using thin-layer chromatography (TLC) (Methods).

POPC seemed to be comparatively more stable (stable up to 30 DH–RH cycles) in nanopure water (Figure 2.7C) than in the Panamic water reaction. TLC band corresponding to POPC (band indicated by black solid arrow) disappeared by 30 DH–RH cycles in the Panamic water sample reaction (Figure 2.7D). With increasing DH–RH cycles, there were also some new bands that were observed, which correspond to the degradation products. These were mainly fatty acids (palmitic acid and oleic acid) and monoacylated phospholipid, whose identity was confirmed by high-resolution mass spectrometry (HRMS). As mentioned earlier, fatty acids (oleic acid) on their own have a tendency to form vesicles at pH 8. So, the vesicles observed in the Panamic water sample after 30 cycles of DH–RH using microscopy (Figure 2.7B) could also be pure fatty acid vesicles or a blended membrane composed of POPC and fatty acid. This is interesting as POPC, being a diacyl structure, can provide stability toward fluctuating pH conditions and the presence of metal ions, while the presence of fatty acid can allow the membrane to exchange material with the bulk solution [48–53]. These are properties that have been argued to be highly desirable in model protocellular membranes.

2.4. Discussion

Cyclic nucleotides have been hypothesized to undergo oligomerization via base-catalyzed reaction [21]. Nonenzymatic oligomerization of purine cyclic nucleotides has been demonstrated under completely dry conditions [15,22]. However, to our knowledge, only a few studies have tried to investigate the nonenzymatic oligomerization of cyclic pyrimidine nucleotides [14,17]. The present study illustrates the enzyme-free oligomerization resulting in the formation of intact informational oligomers in both a purine (cAMP) and a pyrimidine (cCMP) system, under prebiotically relevant DH–RH conditions as well as in a hot spring water sample reminiscent of early Earth environment. It has been reported that short oligomers (up to trimers) can enhance the rate of template-directed replication by about threefold [2–5]. Thus, the intact oligomers reported in the present study can potentially assist in the information propagation in template-directed replication scenarios. Our study demonstrates that prebiotically relevant DH– RH cycles can indeed facilitate the

enzyme-free oligomerization of, both, 2′, 3′ and 3′, 5′ isomers of cyclic nucleotides of both purine and pyrimidine systems. The dry phase facilitates the condensation reaction, although, the rehydration phase is necessary for effective collisions to occur between monomers and/or the growing oligomers. The latter, however, can also enhance the back hydrolysis of oligomers that are already present. Therefore, there is always a dynamical interplay of oligomerization and hydrolysis in these geological settings.
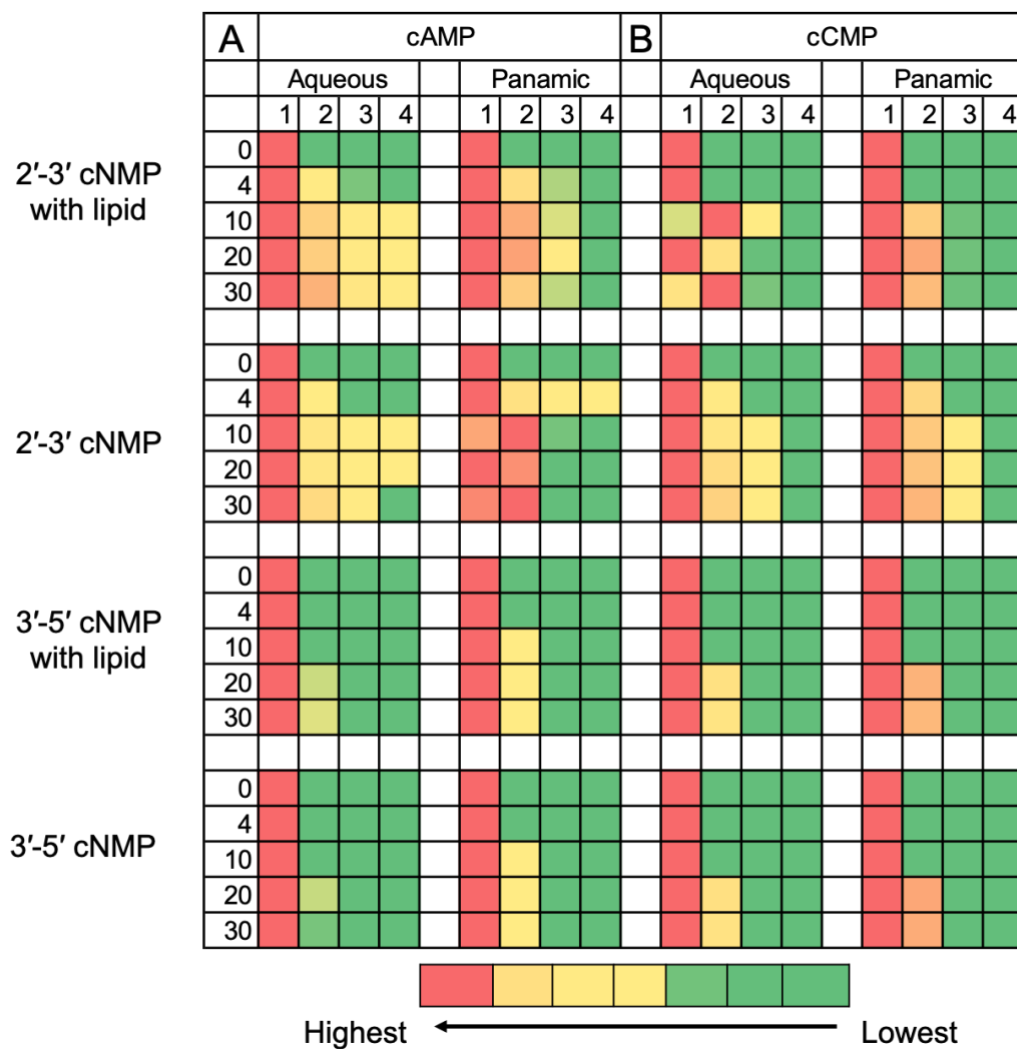
Figure 2.8. Heat map of all the species observed in enzyme-free oligomerization reactions using LC-MS/MS. (A, B) The species were observed in different reactions of cAMP and cCMP, respectively, as mentioned toward the left of the figure. The numbers, viz. 0, 4, 10, 20, and 30, on the left axis of the heat map indicate the number of DH–RH cycles. Lanes 1–4 show the presence of cNMP, dimer, trimer, and tetramer,

respectively. Each block depicts a single reaction with rows and columns, signifying different time points and different species observed, respectively. The numerical data used for this analysis are presented in Table 2.4.

Table 2.4. Nonenzymatic oligomerization of cyclic nucleotides as observed using LC-MS analysis.

		cAMP							
		Nanopure water				Panamic water			
	Cycles	cNMP	Dimer	Trimer	Tetramer	cNMP	Dimer	Trimer	Tetramer
2′–3′ cAMP with lipid	0	1×10^{07}				6×10^{06}			
	4	1×10^{07}	1×10^{05}	6×10^{03}		8×10^{06}	1×10^{06}	2×10^{05}	
	10	5×10^{06}	1×10^{06}	1×10^{05}	5×10^{04}	2×10^{06}	1×10^{06}	8×10^{04}	
	20	3×10^{06}	8×10^{05}	8×10^{04}	2×10^{04}	1×10^{06}	8×10^{05}	8×10^{04}	
	30	2×10^{06}	8×10^{05}	8×10^{04}	2×10^{04}	2×10^{06}	6×10^{05}	6×10^{04}	
2′–3′ cAMP	0	6×10^{06}				1×10^{07}			
	4	1×10^{07}	1×10^{05}			1×10^{07}	9×10^{05}	9×10^{04}	7×10^{04}
	10	1×10^{07}	6×10^{05}	5×10^{04}	5×10^{04}	6×10^{05}	9×10^{05}	4×10^{04}	
	20	1×10^{07}	5×10^{05}	7×10^{04}	6×10^{04}	6×10^{05}	5×10^{05}		
	30	4×10^{06}	6×10^{05}	7×10^{04}		5×10^{05}	6×10^{05}		
3′–5′ cAMP with lipid	0	1×10^{07}				2×10^{07}			
	4	9×10^{06}				1×10^{07}			
	10	1×10^{07}				1×10^{07}	2×10^{04}		
	20	1×10^{07}	3×10^{04}			1×10^{07}	2×10^{04}		
	30	9×10^{06}	3×10^{04}			1×10^{07}	3×10^{04}		
3′–5′ cAMP	0	1×10^{07}				2×10^{07}			
	4	8×10^{06}				1×10^{07}			
	10	8×10^{06}				2×10^{07}	5×10^{03}		
	20	1×10^{07}	2×10^{04}			1×10^{07}	3×10^{03}		
	30	7×10^{06}	4×10^{03}			2×10^{07}	9×10^{03}		

		cCMP							
		Nanopure water				Panamic water			
	Cycles	cNMP	Dimer	Trimer	Tetramer	cNMP	Dimer	Trimer	Tetramer
2′–3′ cCMP with lipid	0	9×10^{06}				8×10^{06}			
	4	1×10^{07}				5×10^{06}	2×10^{06}	1×10^{05}	
	10	1×10^{04}	5×10^{05}	2×10^{04}		4×10^{06}	1×10^{06}	6×10^{04}	
	20	4×10^{06}	7×10^{05}	1×10^{04}		2×10^{06}	1×10^{06}	4×10^{04}	
	30	7×10^{04}	5×10^{05}	5×10^{03}		2×10^{06}	8×10^{05}	2×10^{04}	
2′–3′ cCMP	0	7×10^{06}				6×10^{06}			
	4	8×10^{06}	1×10^{05}	4×10^{03}		6×10^{06}	1×10^{06}	7×10^{04}	
	10	8×10^{06}	4×10^{05}	1×10^{04}		4×10^{06}	1×10^{06}	6×10^{04}	
	20	5×10^{06}	7×10^{05}	1×10^{04}		3×10^{06}	1×10^{06}	4×10^{04}	
	30	4×10^{06}	7×10^{05}	1×10^{04}		3×10^{06}	1×10^{06}	3×10^{04}	
3′–5′ cCMP with lipid	0	2×10^{04}				4×10^{04}			
	4	2×10^{05}				2×10^{05}	6×10^{04}		
	10	1×10^{05}	2×10^{04}			3×10^{05}	1×10^{05}		
	20	2×10^{05}	1×10^{04}			4×10^{05}	2×10^{05}		
	30	4×10^{05}	2×10^{04}			4×10^{05}	2×10^{05}		
3′–5′ ccMP	0	4×10^{04}				6×10^{04}			
	4	1×10^{05}	3×10^{04}			3×10^{05}	3×10^{04}		
	10	1×10^{05}	2×10^{04}			2×10^{05}	7×10^{04}		
	20	1×10^{05}	1×10^{04}			2×10^{05}	8×10^{04}		
	30	4×10^{05}	3×10^{04}			3×10^{05}	1×10^{05}		

The abundance of linear oligomeric species observed in the nonenzymatic oligomerization reactions of cAMP and cCMP in both nanopure water and Panamic water. All the reactions were analyzed using TOF-MS and confirmed with TOF-MS–MS analysis.

Subsequent cycles of DH–RH can act as a kinetic trap by pushing the reaction toward oligomerization, thereby facilitating the formation of longer oligomers over

multiple cycles [12,13]. In these settings, lipids can provide a boundary condition and facilitate the formation of a protocellular entity by encapsulating and thus, conferring protection to select solute molecules (as seen in the case of 2′, 3′ cAMP reactions in both the water samples) [30]. In addition to the formation of intact linear oligomers, the formation of oligomers with cyclic ends (cyclic oligomers) could also potentially be facilitated in the reactions studied. When present in reasonable concentrations, these cyclic oligomers could possibly undergo ligation-like reactions resulting in longer oligomers.

Nonetheless, the structure of these cyclic oligomers cannot be unequivocally ascertained using LC-MS as the mass of the covalently linked cyclic oligomer and noncovalently linked stacked oligomers happen to be the same. This is an aspect that has also been highlighted in other studies too [54]. In the hot spring water sample reactions, the oligomers seemed to form more readily, as was observed, for example, in the reactions involving 2′, 3′ cAMP (tetramer was observed after 4 d of DH–RH cycles) (Figure 2.8A). However, the oligomers formed were also found to be susceptible to hydrolysis, potentially because of the presence of different metal ions or other co-solutes in these scenarios. Interestingly, 2′, 3′ cNMP, upon oligomerization could potentially yield 2′, 5′ as well as 3′, 5′ linked phosphodiester bonds (Figure 2.2) [22]. On the contrary, 3′, 5′ cNMP can oligomerize to yield only 3′, 5′ linked phosphodiester bonds. Given this, it is logical to assume that the enzyme-free oligomerization of both the cyclic nucleotide isomers (2′, 3′ and 3′, 5′) might result in products with mixed phosphodiester linkages. Significantly, 2′, 5′ phosphodiester linkages are more prone to hydrolysis even at pH8 and 40 C with a rate of hydrolysis of about 0.01/h as compared to 0.003/h for 3′, 5′ linkages [55], indicating that this could only get even more exaggerated at 90 C (that we used in our experiments). Our study emphasizes the fact that even though the speed of oligomerization of 3′, 5′ cNMPs is compromised, the resultant oligomers are found to be stable even after 30 DH–RH cycles (Figure 2.8). The comparatively low stable products resulting in 2′, 3′ cNMP oligomerization (especially in cAMP reactions in Panamic water) could be due to the formation of more hydrolyzable 2′, 5′ linked oligomers, in comparison to the 3′, 5′ linked oligomers. The results from our reactions pertaining to the hot spring water

sample (a proxy for prebiotically pertinent conditions), could, therefore, be indicative for how the selection of 3′, 5′ phosphodiester linkages might have been driven by such relevant selection pressures. We propose that in prebiotic environments (e.g., hot spring waters) or in the presence of metal ions, selection for the 3′, 5′ phosphodiester linkage over time might have been naturally facilitated. Additionally, our results seemed to indicate that the presence of ions and other co-solutes (such as silicates) in the hot spring water sample, could effectively act as a selection pressure playing an important role in shaping the evolutionary landscape of a putative RNA World. Altogether, this study provides a comprehensive and robust means by which enzyme-free oligomerization of intrinsically activated, prebiotically pertinent nucleotides might have occurred in a heterogenous prebiotic soup.

2.5. References

1. Gilbert, W. Origin of life: The RNA world. *Nature* **319**, (1986).
2. Kaiser, A. & Richert, C. Nucleotide-Based Copying of Nucleic Acid Sequences without Enzymes. *J. Org. Chem.* **78**, 793–799 (2013).
3. Li, L. *et al.* Enhanced Nonenzymatic RNA Copying with 2-Aminoimidazole Activated Nucleotides. *J. Am. Chem. Soc.* **139**, 1810–1813 (2017).
4. Tam, C. P. *et al.* Downstream oligonucleotides strongly enhance the affinity of GMP to RNA primer-template complexes. *J. Am. Chem. Soc.* **139**, 571–574 (2017).
5. O'Flaherty, D. K. *et al.* Copying of Mixed-Sequence RNA Templates inside Model Protocells. *J. Am. Chem. Soc.* **140**, 5171–5178 (2018).
6. Puthenvedu, D., Janas, T., Majerfeld, I. & Yarus, M. Poly (U) RNA-templated synthesis of AppA. *RNA* 1–8 (2015). doi:10.1261/rna.052696.115.
7. Puthenvedu, D., Majerfeld, I. & Yarus, M. Non-Watson-Crick RNA synthesis suited to origin functions. *RNA* **24**, 90–97 (2017).
8. Majerfeld, I., Puthenvedu, D. & Yarus, M. Cross-backbone templating ; ribodinucleotides made on poly (C). *RNA* 397–407 (2016). doi:10.1261/rna.054866.115.

9.   Lohrmann, R. & Orgel, L. E. Prebiotic Activation Processes. *Nature* **244**, 418–419 (1973).

10.   Blain, J. C. & Szostak, J. W. Progress Toward Synthetic Cells. *Annu. Rev. Biochem.* **83**, 615–640 (2014).

11.   Yi, R., Hongo, Y. & Fahrenbach, A. C. Synthesis of imidazole-activated ribonucleotides using cyanogen chloride. *Chem. Commun.* 1–4 (2018). doi:10.1039/C7CC08489G

12.   Rajamani, S. *et al.* Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Orig. Life Evol. Biosph.* **38**, 57–74 (2008).

13.   Mungi, C. V., Rajamani, S., Mungi, C. V. & Rajamani, S. Characterization of RNA-Like Oligomers from Lipid-Assisted Nonenzymatic Synthesis: Implications for Origin of Informational Molecules on Early Earth. *Life (Basel, Switzerland)* **5**, 65–84 (2015).

14.   Tapiero, C. M. & Nagyvary, J. Prebiotic Formation of Cytidine N ucleotides. *Nature* **231**, 42–43 (1971).

15.   Verlander, M. S., Lohrmann, R. & Orgel, L. E. Catalysts for the self-polymerization of adenosine cyclic 2',3'-phosphate. *J. Mol. Evol.* **2**, 303–316 (1973).

16.   Verlander, M. S. & Orgel, L. E. Analysis of High Molecular-Weight Material From Polymerization of Adenosine Cyclic 2',3'-Phosphate. *J. Mol. Evol.* **3**, 115–120 (1974).

17.   Costanzo, G. *et al.* Nonenzymatic Oligomerization of 3′,5′-Cyclic CMP Induced by Proton and UV Irradiation Hints at a Nonfastidious Origin of RNA. *ChemBioChem* **18**, 1535–1543 (2017).

18.   Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).

19.   Powner, M. W., Sutherland, J. D., Szostak, J. W. & Matthew W. Powner,*,† John D. Sutherland, ‡ and Jack W. Szostak†. Chemoselective multicomponent one-pot assembly of purine precursors in water. *J. Am. Chem. Soc.* **132**, 16677–16688 (2010).

20.   Suárez-Marina, I. *et al.* Integrated synthesis of nucleotide and nucleosides influenced by amino acids. *Commun. Chem.* **2**, 1–8 (2019).

21. Costanzo, G. *et al.* Generation of RNA Molecules by a Base-Catalysed Click- Like Reaction. 999–1008 (2012). doi:10.1002/cbic.201200068

22. Costanzo, G. *et al.* Non-Enzymatic Oligomerization of 3', 5' Cyclic AMP. *PLoS One* **11**, (2016).

23. Higgs, P. G. The effect of limited diffusion and wet–dry cycling on reversible polymerization reactions: Implications for prebiotic synthesis of nucleic acids. *Life* **6**, 1–16 (2016).

24. Ross, D. & Deamer, D. Dry/Wet Cycling and the Thermodynamics and Kinetics of Prebiotic Polymer Synthesis. *Life* **6**, 28 (2016).

25. Becker, S. *et al.* Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat. Commun.* **9**, 1–9 (2018).

26. Becker, S. *et al.* Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science (80-.).* **82**, 76–82 (2019).

27. Mamajanov, I. *et al.* Ester formation and hydrolysis during wet-dry cycles: Generation of far-from-equilibrium polymers in a model prebiotic reaction. *Macromolecules* **47**, 1334–1343 (2014).

28. Forsythe, J. G. *et al.* Ester-Mediated Amide Bond Formation Driven by Wet-Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. *Angew. Chemie - Int. Ed.* **54**, 9871–9875 (2015).

29. Toppozini, L., Dies, H., Deamer, D. W. & Rheinstädter, M. C. Adenosine Monophosphate Forms Ordered Arrays in Multilamellar Lipid Matrices: Insights into Assembly of Nucleic Acid for Primitive Life. *PLoS One* **8**, 62810 (2013).

30. Damer, B. & Deamer, D. Coupled Phases and Combinatorial Selection in Fluctuating Hydrothermal Pools: A Scenario to Guide Experimental Approaches to the Origin of Cellular Life. *Life* **5**, 872–887 (2015).

31. Giovanna, C., Pino, S., Ciciriello, F. & Di Mauro, E. Generation of long RNA chains in water. *J. Biol. Chem.* **284**, 33206–33216 (2009).

32. Deamer, D., Singaram, S., Rajamani, S., Kompanichenko, V. & Guggenheim, S. Self-assembly processes in the prebiotic environment. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361**, 1809–18 (2006).

33. Joshi, M. P., Samanta, A., Tripathy, G. R. & Rajamani, S. Formation and Stability

of Prebiotically Relevant Vesicular Systems in Terrestrial Geothermal Environments. *Life* **7**, 51 (2017).

34. Khorana, H. G., Tener, G. M., Wright, R. S. & Moffatt, J. G. Cyclic Phosphates. III. Some General Observations on the Formation and Properties of Five-, Six- and Seven-membered Cyclic Phosphate Esters. *J. Am. Chem. Soc.* **79**, 430–436 (1957).

35. Gerlt, J. A., Westheimer, P. H. & Sturtevant, J. H. *The Enthalpies of Hydrolysis of Acyclic, Monocyclic, and Glycoside Cyclic Phosphate Diesters**. *The Journal of biological chemistry* **250**, (1974).

36. Deamer, D. W. & Barchfeld, G. L. Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions. *J. Mol. Evol.* **18**, 203–206 (1982).

37. Black, R. A. *et al.* Nucleobases bind to and stabilize aggregates of a prebiotic amphiphile, providing a viable mechanism for the emergence of protocells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13272–6 (2013).

38. Sasidharan, S. *et al.* Interaction of the mononucleotide UMP with a fluid phospholipid bilayer †. *Soft Matter* **15**, 8129–8136 (2019).

39. Chen, I. A., Salehi-Ashtiani, K. & Szostak, J. W. RNA catalysis in model protocell vesicles. *J. Am. Chem. Soc.* **127**, 13213–13219 (2005).

40. Chen, I. A. & Walde, P. From self-assembled vesicles to protocells. *Cold Spring Harb. Perspect. Biol.* **2**, 1–14 (2010).

41. Deguzman, V., Vercoutere, W., Shenasa, H. & Deamer, D. Generation of oligonucleotides under hydrothermal conditions by non-enzymatic polymerization. *J. Mol. Evol.* **78**, 251–262 (2014).

42. Lailach, G., Thompson, T., Miner, G. B.-C. C. & 1968, U. Absorption of pyrimidines, purines, and nucleosides by Co-, Ni-, Cu-, and Fe (III)-montmorillonite (Clay-organic studies XIII). *Clays Clay Miner.* **16**, 295–301 (1968).

43. Ferris, J. P. Mineral Calalysis and Prebiotic Synthesis : Montmorillonite-Catalysed Formation of RNA. *Elements* **1**, 145–149 (2005).

44. Joshi, P. C., Aldersley, M. F., Zagorevskii, D. V. & Ferris, J. P. A Nucleotide Dimer Synthesis Without Protecting Groups Using Montmorillonite as Catalyst. *Nucleosides, Nucleotides and Nucleic Acids* **31**, 536–566 (2012).

45. Blasko, A. & Bruice, T. C. Recent Studies of Nucleophilic, General-Acid, and Metal Ion Catalysis of Phosphate Diester Hydrolysis. *Acc. Chem. Res.* **32**, 475–484 (1999).

46. Pandey, S. *et al.* Ladakh: Diverse, high-altitude extreme environments for off-earth analogue and astrobiology research. *Int. J. Astrobiol.* 1–21 (2019). doi:10.1017/S1473550419000119

47. Monnard, P.-A. & Walde, P. Current ideas about prebiological compartmentalization. *Life* **5**, 1239–1263 (2015).

48. Luisi, P. L., Walde, P. & Oberholzer, T. Lipid vesicles as possible intermediates in the origin of life. *Curr. Opin. Colloid Interface Sci.* **4**, 33–39 (1999).

49. Mansy, S. S. Model protocells from single-chain lipids. *Int. J. Mol. Sci.* **10**, 835–843 (2009).

50. Budin, I. & Szostak, J. W. Physical effects underlying the transition from primitive to modern cell membranes. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5249–5254 (2011).

51. Dalai, P., Ustriyana, P. & Sahai, N. Aqueous magnesium as an environmental selection pressure in the evolution of phospholipid membranes on early earth. *Geochim. Cosmochim. Acta* **223**, 216–228 (2018).

52. Jin, L., Kamat, N. P., Jena, S. & Szostak, J. W. Fatty Acid/Phospholipid Blended Membranes: A Potential Intermediate State in Protocellular Evolution. *Small* **1704077**, 1–9 (2018).

53. Sarkar, S., Dagar, S., Verma, A. & Rajamani, S. Compositional heterogeneity confers selective advantage to model protocellular membranes during the origins of cellular life. *Sci. Rep.* **10**, 4483 (2020).

54. Kaddour, H. *et al.* Nonenzymatic RNA Oligomerization at the Mineral−Water Interface: An Insight into the Adsorption−Polymerization Relationship. *J. Phys. Chem. C* **122**, 29386–29397 (2018).

55. D. A. Usher & McHale, A. H. Hydrolytic stability of helical RNA: A selective advantage for the natural 3',5'- bond. *Proc Natl Acad Sci U S A* **250**, 5059–5067 (1976).

Chapter 3

Nonenzymatic template-directed primer extension using 2′-3′ cyclic nucleotides under wet-dry cycles
(Adapted from Dagar et. al 2022; bioRxiv)

3.1. Introduction

Contemporary biology relies on dedicated biopolymers such as DNA/RNA to store and propagate genetic information, and proteins that enable the catalysis of reactions [1]. As primitive cells are considered to be devoid of a refined and evolved molecular machinery, these functions might have been carried out by prebiotically relevant, simple molecular systems. Towards this, the well-explored 'RNA World Hypothesis' suggests RNA as having been the first biopolymer to have emerged due to its ability to serve both as a catalyst and a genetic material [2]. In the absence of enzymes, the formation and replication of RNA would have had to be nonenzymatic; driven by both the monomeric species involved and the environmental conditions. Previous attempts to demonstrate the enzyme-free primer extension of RNA either involved chemically modified primer (3′-dideoxy-NH_2 modified primer (NH_2-primer)), nucleotides (imidazole, methylimidazole or aminoimidazole at 5′ end of the nucleotide), or both, and were performed at ambient temperature (24°C) [3–11]. Although some studies have shown a plausible route for the formation of such activated nucleotides under prebiotic settings [12–16], their presence in significant amounts on the early Earth is questionable. Pertinently, their susceptibility towards hydrolysis at high temperature and in aqueous conditions, renders them as unsuitable substrates, putting into question their use as prebiotically compatible reactants. Pertinently, other relevant studies have investigated the possibility of in situ activation in primer extension reactions using high concentrations (0.1-0.8 M) of condensing agents such as 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), 1-methyladenine and 1-ethylimidazole (1-EtIm) etc. at cold to ambient temperatures (0°C to 20°C)[17–19]. However, primer extension reactions with non-activated nucleotides i.e., nucleoside 5′-monophosphate (5′-NMP) without the use of any activating/condensing agent under simulated early Earth conditions (high temperature) are limited. In 2018, Bapat *et. al.* investigated phospholipid-assisted template-directed NH_2-primer extension reactions using non-activated nucleotides i.e., nucleoside 5′-monophosphate (5′-NMP), under acidic hydrothermal conditions [20]. The presence of amino group in the primer

makes it a better nucleophile as compared to hydroxyl-terminated primer, facilitating its attack on the phosphorus group of the incoming nucleotide, resulting in the formation of a phosphoramidate bond. These reactions were performed under alternating dehydration-rehydration (DH-RH) cycles at high temperature (90°C) and acidic pH (pH 2). DH-RH cycles are a recurring geological phenomenon facilitated readily in terrestrial hydrothermal pools and have been demonstrated to facilitate condensation reactions and the formation of different biopolymers under prebiotic settings [21,22]. The dry phase concentrates the substrate molecules and the rehydrated phase facilitates mixing, thus, increasing the chances of effective collisions and redistribution of the growing oligomers. Lipids have been shown to form multilamellar films during the dry phase, which upon rehydration can encapsulate the surrounding solutes (e.g. informational oligomers), resulting in protocell-like entities [23–25]. However, similar to the oligomerization reactions, the systematic characterization of the primer extension product obtained under these reaction conditions (high temperature and pH 2), showed depurination of potentially the incoming nucleotide that resulted in loss of the information moiety [20,26,27]. Importantly, similar nonenzymatic template-directed primer extension reactions with 5′-NMPs and hydroxyl-terminated primer (OH-primer), resulted in no prominent primer extension [20]. This emphasizes the necessity of using prebiotically relevant alternative monomers, which do not require acidic pH for the formation of a phosphodiester bond, thereby alleviating the problem of the loss of information moiety.

Recently, we demonstrated the nonenzymatic oligomerization of prebiotically relevant cyclic nucleoside monophosphates (cNMPs) under DH-RH cycles that yielded intact RNA oligomers [28]. Their prebiotically possible synthesis, stability to high temperatures, and being intrinsically active due to the presence of an intramolecular cyclic group, make cNMPs a potential substrate for nonenzymatic oligomerization and replication reactions [12,13,29–34]. However, to our knowledge, no study has been reported wherein template-directed primer extension has been attempted using cNMPs. Therefore, we investigated the use of cNMPs in template-directed primer extension reactions of RNA. The two isomers of cNMP (2′-3′cNMP or 3′-5′cNMP) are known to possess different

stability towards hydrolysis, and hence differ in their reactivity and oligomerization potential [28]. Given this, the template-directed primer extension reactions were performed with both the structural isomers i.e., 2′-3′ cNMP and 3′-5′ cNMP using a purine, (cAMP) or a pyrimidine (cCMP) as the incoming monomer. The reactions were undertaken using both the OH-primer and NH2-primer (latter has been used extensively in many previous studies). As lipids have been shown to impart protection to RNA against hydrolysis under DH-RH conditions [20,26,35], the effect of the presence of lipids was also evaluated in these reactions. Previous studies that evaluated lipid-assisted template-directed primer extension have predominantly employed pure phospholipid (PL) membranes [20,36]. Although insightful, such membranes are impermeable to polar molecules and would require complex membrane proteins to facilitate such processes [37]. The exchange of material with the surroundings is an essential feature of robust protocells, thus, making it implausible for just pure PL membranes to act as prebiotic compartments. In this regard, few other studies have investigated fatty acid/PL blended (hybrid) membranes as model protocellular compartments [38]. Presence of a single chain amphiphile (SCA) like fatty acids in these blended membranes, enhanced the permeability of these compartments, while the use of a diacyl chain amphiphile like PL, enabled maintaining their stability against environmental fluctuations. Given this, we evaluated the effect of lipids on nonenzymatic primer extension reactions using three different model protocellular membranes i.e., only palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), and two binary systems of POPC with glycerol 1-monooleate (GMO) and POPC with oleic acid; both present in 1:1 molar ratio in the binary systems.

Our results showed the extension of the primer with two intact nucleotide additions, in the reactions involving 2′-3′ cNMPs at pH 8 and under DH-RH conditions. No significant primer extension was observed in the reactions involving 3′-5′ cNMPs. Importantly, the primer extension by two intact nucleotides was observed for both the OH-primer and NH2-primer, which would have resulted in phosphodiester and phosphoramidate linkage, respectively (Scheme 3.1). The primer extension was observed with both 2′-3′ cAMP as well as 2′-3′ cCMP. Pertinently, in the 2′-3′ cCMP reactions, the presence of

lipids was observed to significantly increase the yield of the extended product. Altogether, this study provides a comprehensive approach by which enzyme-free information propagation might have been facilitated on the early Earth in a putative RNA World that involves the use of prebiotically realistic, 'intrinsically active' monomers.
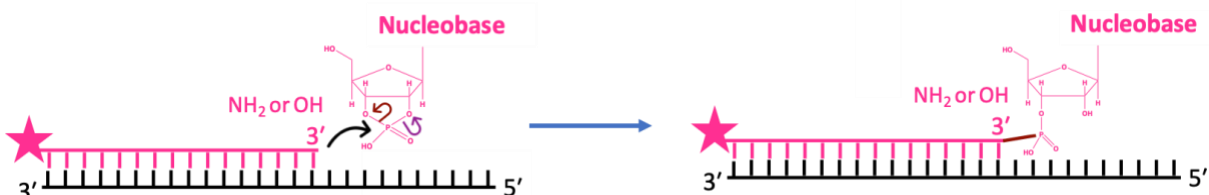
Scheme 3.1: Scheme showing template-directed primer extension using cyclic nucleotide and either NH$_2$-primer or OH-primer, resulting in phosphoramidate bond or phosphodiester bond, respectively.

3.2. Materials and Methods

3.2.1. Materials

The monosodium salts of all four cyclic monophosphates viz. Adenosine 3′, 5′ cyclic monophosphate (3′, 5′ cAMP), Adenosine 2′, 3′ cyclic monophosphate (2′, 3′ cAMP), Cytidine 3′, 5′ cyclic monophosphate (3′, 5′ cCMP) and Cytidine 2′, 3′ cyclic monophosphate (2′, 3′ cCMP) were purchased from Sigma-Aldrich (Bangalore, India) and used without further purification. 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) was purchased from Avanti Polar Lipids Inc (Alabaster, AL, USA); other co-surfactants i.e., glycerol 1-monooleate (GMO) and oleic acid (OA) were purchased from Nu-Chek-Prep (Elysian, MN, USA) and all were used without further purification. All other reagents used were of analytical grade and purchased from Sigma-Aldrich (Bangalore, India). The RNA primers used in this study are Amino-G (NH$_2$-primer; acquired from Keck laboratory, Yale, USA), Hydroxyl-G (OH-primer) and 10-mer hydroxyl-G (10-mer OH-primer) (acquired from Thermo Fisher Scientific, USA) primers. Both Amino-G and Hydroxyl-G primers were fluorescently labelled with Cyanine 3 (Cy3) on their 5′-end for facilitating their detection on polyacrylamide gel electrophoresis

(PAGE). The Amino-G primer terminates with a 3′-amino-2′, 3′-dideoxynucleotide (Metkinen, Finland) while the Hydroxyl-G primer terminates with a canonical ribonucleotide. The sequences of the primers and templates are as given below, with the template base indicated in bold:

Primer Amino-G (Amino/ NH_2-primer): 5′ GG GAU UAA UAC GAC UCA CUG_{NH2}-3′

Primer Hydroxyl-G (Hydroxyl/ OH-primer): 5′ GG GAU UAA UAC GAC UCA CUG_{OH}-3′

10-mer primer hydroxyl-G (10-mer OH-primer): 5′ C GAC UCA CUG_{OH}-3′

Template U: 5′ AGU GAU CU**U** CAG UGA GUC GUA UUA AUC CC_{OH}-3′

Template G: 5′ AGU GAU CU**G** CAG UGA GUC GUA UUA AUC CC_{OH}-3′

3.2.2. Methods

3.2.2.1. Reaction setup

In a typical reaction, 2.5 µM of the template (U in the case of cAMP and G in the case of cCMP) was annealed with 1.25 µM of the Cy3 labelled primer (either OH-primer or NH_2-primer), by heating them at 95 °C for 5 minutes and then cooling to room temperature (RT). 5 mM of the corresponding cyclic nucleotide monophosphate (cNMP) was added to the annealed primer-template complex. 1 mM of either pure POPC or binary suspension of POPC:GMO::1:1 or POPC:OA::1:1 was added in case of the lipid-assisted reactions. The pH of the mixture was adjusted to ~ 8, unless specified. The reaction mixtures were then subjected to repeated cycles of Dehydration-Rehydration (DH-RH) at 90°C, using 24 hours per cycle. After each dehydration cycle, the reaction mixture was rehydrated with 1 mM of the respective cNMP to compensate for potential loss of substrate. The samples were withdrawn after the rehydration step for analysis after regular intervals (i.e., starting of the reaction, after one day, two days and five days, respectively). These samples were then analyzed using denaturing PAGE.

3.2.2.2. PAGE analysis of the reaction products

After varying time periods (i.e., starting of reaction, after one day, two days and five days), the samples were collected in 1X TBE buffer containing 8 M urea. The sample

volumes withdrawn at different time points were adjusted to a standardized amount to compensate for the degradation of RNA primer that occurs over multiple cycles of DH-RH. A non-fluorescent competitor RNA (untagged primer), whose sequence was exactly the same as that of the tagged primer, was added in at least 10 times excess as compared to the fluorescently tagged primer, in all the samples that were to be run on the PAGE gel. This was done in order to successfully separate the fluorescent primer and the extended product from the template for unhindered and effective gel analysis. The reaction products were analyzed on 20% denaturing PAGE. The gels were then scanned with an Amersham Typhoon Biomolecular imager (GE Healthcare) at 550 PMT and 100-micron resolution setting, using the Cy3 (532 nm) excitation laser. The gel images were subsequently processed using ImageQuant v8.2 software for quantification of the relevant bands.

3.2.2.3. Quantification of the reaction products

The yield (%) of the extended products of the primer was calculated by using the following formula:

$$Yield\ (\%)\ of\ the\ extended\ product = \left(\frac{\left(\frac{In}{I0} \right) * 100}{Vn0} \right) \qquad \text{Eq. 1}$$

Where, I_n and I_0 are the intensities of the band at the nth cycle and the 0th cycle, respectively, and V_{n0} is the volume fold change in n^{th} cycle with respect to 0th cycle, calculated by:

$$Vn0 = \frac{Vn}{V0} \qquad \text{Eq. 2}$$

3.2.2.4. LC-MS analysis of the extended products

For LC-MS analysis, primer extension reactions were performed using the 10-mer OH-primer. This was done in order to detect the undegraded extended primer under the conditions optimal for LC-MS analysis, as the longer the RNA, the higher is the possibility of its degradation upon ionization. A mixture of 1.25 µM of 10-mer primer and 5 mM of 2′, 3′ cNMP was subjected to DH-RH cycles at pH 8. After each dehydration cycle (24 hours per cycle), the reaction mixture was rehydrated with 1 mM of the respective cNMP. After two DH-RH cycles, all the reaction mixture timepoints from a

specific reaction were pooled together (total volume 5 ml), in order to increase the total extended product to be present in detectable amount. This mixture was then lyophilized and resuspended in 52 µl, out of which 50 µl was loaded onto LC-MS. The reaction mixtures were separated on a Zorbax C8 column (dimensions: 4.6 X 150 mm, particle size: 3.6 µm) (Thermo Scientific), fitted with a guard column. The species were separated on LC using a gradient of solvent A of 95:5 (vol/vol) H_2O/methanol (MeOH) + 0.1% ammonium hydroxide, and solvent B of 60:35:5 (vol/vol) isopropanol/MeOH/H_2O + 0.1% ammonium hydroxide. The gradient involved an isocratic phase of solvent A for 10 minutes, followed by an increase to 100% of solvent B for six minutes and subsequent equilibration with solvent A for seven minutes. Mass spectra of the reaction samples was recorded on a Sciex X500R QTOF mass spectrometer (MS) fitted with an Exion-LC series UPLC (Sciex, CA, USA) using information-dependent acquisition (IDA) scanning method. The acquired data was analyzed using the Sciex OS software. All the mass acquisition was performed using Electron spray ionization (ESI) in the negative mode with the following parameters: turbo spray ion source, medium collision gas, curtain gas = 30 L/min, ion spray voltage = -4500 V (negative mode), at 500 °C. TOF-MS acquisition was done using a declustering potential of -80 V, and -10 V collision energy. As the mass acquisition was carried out in the negative mode, the observed masses correspond to the mass of the nH^- (where n is the number of deprotonation sites) adducts of the parent molecule. The presence of a specific species/molecule was confirmed by the presence of precursor mass within 10 ppm error.

3.3. Results

3.3.1. Primer extension using NH₂-primer and cyclic nucleotides under DH-RH conditions.

Given that cNMPs are prebiotically relevant nucleotides and can oligomerize nonenzymatically [28,30,31,39,40], we asked whether these can act as a substrate for enzyme-free RNA copying. Firstly, we investigated template-directed primer extension using an amino terminated primer (NH_2-primer), with 2′, 3′ cAMP as the incoming nucleotide, under multiple DH-RH cycles at 90°C and varying pH (2, 8 and 10) (Figure 3.1a). As alluded to earlier, the rehydrated reaction mixture was withdrawn after various time periods and analyzed using denaturing PAGE (detailed in the Methods section). A reaction with known single nucleotide addition to the original 20-mer amino primer was used as the control (indicated by 'N+1' lane; 'N' indicates 20-mer amino primer and 'N+1' indicates the extended 21-mer product).

At pH 2, the intensity of 'N' band decreased with increasing DH-RH cycles, with no extension observed. This could be because of low stability of RNA under these harsh conditions (pH 2 and 90°C). However, at both pH 8 and 10, two new bands above the 'N' band, indicating the presence of species with higher molecular weight, were observed (Figure 3.1a). These new bands indicated the possibility of extension by one (N+1) and two nucleotides (N+2), respectively (Figure 3.1a). In the case of pH 10, up to 9.4% of the extended product was observed after one DH-RH cycle. Nonetheless, the intensity of the bands corresponding to the extended products was observed to decrease significantly with increasing number of DH-RH cycles. Only 5.6 % of the extended product was observed after two cycles at pH 10, which further decreased to 3.8% after five DH-RH cycles. The reason behind this could be the instability of particularly the phosphodiester bond at alkaline pH. The template-directed reactions carried out at pH 8 gave optimal yield, with up to 15.9 % of primer being extended within one DH-RH cycle (Figure 3.1c). The yield of the extended product was observed to decrease to 12.7 % and 11.9 % after two and five DH-RH cycles, respectively. Pertinently, this extended product was also observed in untemplated reactions (Figure 3.1a). In the case of untemplated reactions, up to 12.4 % primer was observed to extend within one DH-RH cycle. Similar to the template-directed reactions, the yield of extended primer decreased to 10.6 % and 7.7 % after two and five DH-RH cycles, respectively. Given the aforementioned observations, all subsequent reactions were performed at pH 8.

As mentioned earlier, the two cyclic isomers possess different reactivity owing to their corresponding ring strains. Encouraged by the primer extension using 2′, 3′ cAMP, we next investigated the template-directed primer extension reaction using 3′, 5′ cAMP. All other reaction conditions (such as temperature, pH and DH-RH duration) were kept the same as that of the reactions that were done using 2′, 3′ cAMP. 1 mM 3′, 5′ cAMP was used as the rehydrating agent in these reactions. Given the intrinsic higher stability of 3′, 5′ cAMP (over 2′, 3′ cAMP), the reaction mixtures were subjected to ten DH-RH cycles to characterize observable reaction changes.

As observed in Figure 3.1b, with increasing number of DH-RH cycles, the intensity of the 'N' band decreased. However, no clear primer extension bands were observed even after ten DH-RH cycles (Figure 3.1b). This could be potentially because of the higher ring stability of 3′, 5′ cAMP, which makes it comparatively less reactive. Moreover, our reaction conditions of high temperature (90 °C) and alkaline pH (8) results in relatively greater RNA primer degradation when experienced for longer periods (e.g., ten days), as is seen in Figure 3.1b. This further adds to the incompatibility of using 3′, 5′ cAMP, as it potentially requires longer duration to react and facilitate template-directed primer extension reactions.

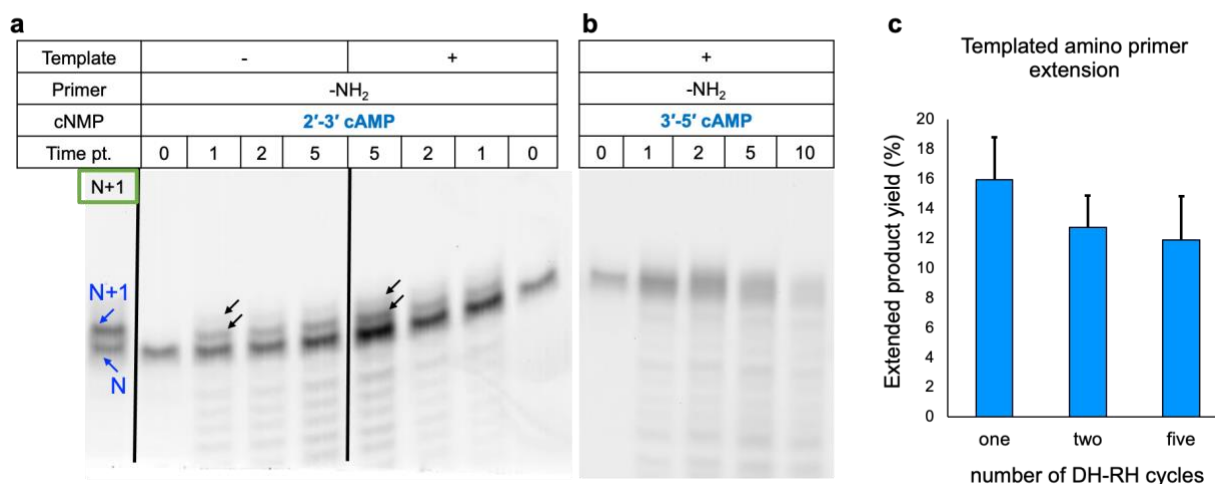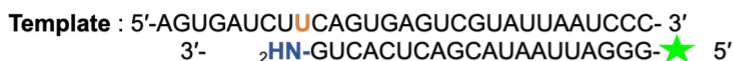Template : 5′-AGUGAUCUUCAGUGAGUCGUAUUAAUCCC- 3′
3′-   2HN-GUCACUCAGCAUAAUUAGGG-⭐ 5′

Figure 3.1: Extension of the NH₂-primer using cAMP as the monomer over multiple cycles of DH-RH. The reactions were performed using 2′, 3′ cAMP (a) or 3′,

5′ cAMP (b), either in the absence (1a, left panel) or presence of template U (1a right panel and b), and analyzed after varying number of DH-RH cycles (cycle 0, cycle 1, cycle 2, cycle 5 and cycles 10). In the N+1 lane, blue arrows indicating 'N' and 'N+1' denote the control 20-mer RNA primer and the primer extended by one nucleotide (21-mer RNA), respectively. The black arrows indicate the extended products in the respective reactions showing up to two nucleotide additions as against the control lane. The black vertical lines have been used to demarcate two reaction sets that were run on the same gel. (c) Yield of the extended primer product in the reactions involving template-directed NH_2-primer reaction, by analyzing denaturing gels using Image Quant v8.2. Y-axis shows the quantified yield percent of the total extended product (sum of both the extended bands), after different DH-RH cycles (viz., 1, 2 and 5) as indicated in the X-axis. The difference was found to be insignificant based on a two-tailed t-test. Error bars = s.d., N = 3.

3.3.2. Primer extension using OH-primer and cyclic nucleotides under DH-RH conditions

After undertaking preliminary template-directed primer extension reactions using NH_2-primer where the extended product is linked with a phosphoramidate bond, we also used OH-primer to check its propensity for extension using cNMPs. It is pertinent to mention that earlier attempts involving template-directed OH-primer extension using non-activated nucleotides, did not result in intact single nucleotide-based extensions [20]. This could be because of the lesser nucleophilicity of the hydroxyl group when compared to an amino group. Nonetheless, we wanted to investigate the capability of cyclic nucleotides to extend OH-primer, which would result in the contemporary phosphodiester inter-nucleotide linkage. Towards this, reactions were performed using OH-primer and 2′, 3′ cAMP. All other reaction conditions (concentrations of reactants, pH, temperature and DH-RH duration) were kept the same as in the reactions that involved the use of NH_2-primer.

Contrary to when non-activated nucleotides ((5′-NMP) were used, two extension bands of the OH-primer were observed in both the untemplated as well as the

template-directed reactions; similar to what was seen with the NH$_2$-primer (Figure 3.2a) reactions. Up to 12.9% of the extension product yield was observed after one DH-RH cycle in these reactions. Up to 12.6% and 11% of extension products persisted even after two and five DH-RH cycles, respectively (Figure 3.2b). In the case of untemplated OH-primer extension reactions, the extended product yield was observed to be 8% within one DH-RH cycle. Upon comparison, up to 11% of the extended product was observed to persist in the presence of template even after five DH-RH cycles, as compared to only ~5.3 % in the untemplated reactions. Upon comparison of the OH-primer reactions with NH$_2$-primer reactions, the difference in the yield of the extended product was observed to be insignificant, irrespective of the presence or absence of the template. This was really encouraging as the extension with OH-primer results in a canonical phosphodiester linkage instead of phosphoramidate linkage (as is the case when using NH$_2$-primer). Moreover, to our knowledge, this is the only reaction to date that has been shown to extend contemporary hydroxyl RNA primer, without the requirement of activating either the nucleotide or the primer. This is exciting as the use of cNMPs would have readily facilitated uphill oligomerization and templated primer extension reactions, without the need for invoking any activation chemistries in the prebiotic scenario. Additionally, similar to the reactions with NH$_2$-primer, no extension bands were observed when 3′, 5′ cAMP was used instead of 2′, 3′ cAMP.

In order to confirm whether the extended products seen in the aforementioned contexts are indeed intact nucleotide additions, we performed the same reaction but with a 10-mer OH-primer (detailed in Methods section), to facilitate a comprehensive analysis of the reaction mixture using LC-MS (detailed in the Methods section). The presence of m/z corresponding to intact nucleotide additions to yield 'N+1' (Figure 3.2c) and 'N+2' (Table 3.1) within ± 5 ppm error, confirmed beyond doubt that the primer extensions in reactions using 2′, 3′ cAMP indeed are extended products with intact informational moieties. This is contrary to scenarios that led to the sugar-phosphate backbone extension but resulted in abasic sites that were a consequence of deglycosylation of the informational entities, resulting in molecules that would lack the capacity for hydrogen bonding-based functionalities [20,26].

Template : 5′-AGUGAUCU**U**CAGUGAGUCGUAUUAAUCCC- 3′
3′- **HO**-GUCACUCAGCAUAAUUAGGG-★ 5′
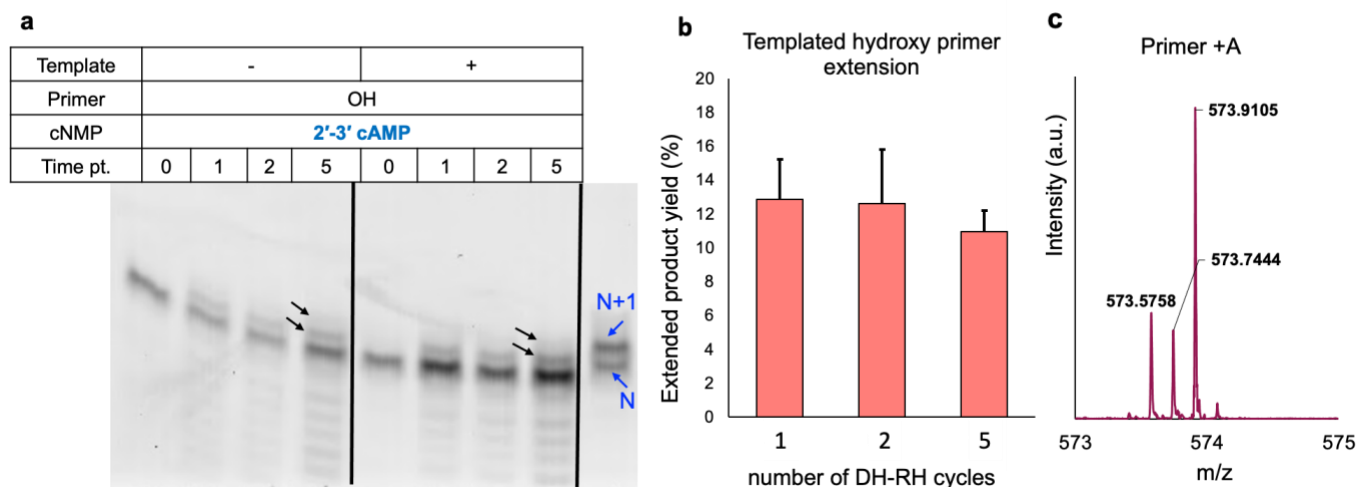
Figure 3.2: Extension of the OH-primer using 2′, 3′ cAMP as the monomer over repeated cycles of DH-RH. (a) The reactions were performed either in the absence (2a, left panel) or presence of template U (2a, right panel) and analyzed after varying number of DH-RH cycles (cycle 0, cycle 1, cycle 2 and cycle 5). In the N+1 lane, blue arrows indicating 'N' and 'N+1' denote the control 20-mer RNA primer and the extended 21-mer RNA, respectively. The black arrows indicate the extended product in the respective reaction mixtures. The black vertical lines have been used to demarcate two reaction sets that were run on the same gel. (b) Yield of the extended primer product with the OH-primer, obtained by analyzing the denaturing gels using Image Quant v8.2. Y-axis shows the quantified yield percent of the total extended product (sum of both the extended bands), after different DH-RH cycles (viz., one, two and five) as is indicated in the X-axis. The difference was found to be insignificant based on a two-tailed t-test. Error bars = s.d., N = 3. (c) Representative TOF-MS spectrum in ESI negative mode for cAMP reactions showing the extended 10-mer primer with the addition of AMP, resulting in a 11-mer primer [M] with m/z corresponding to $[M-6H]^{6-}$ = 573.9105.

Table 3.1: Masses observed during LC-MS characterization of primer extension reactions using 2′, 3′ cAMP, along with calculated parts per million (ppm) error. Masses for one nucleotide and two

nucleotide extensions on to the 10-mer OH-primer have been indicated as 'N+1' and 'N+2', respectively.

2'-3' cAMP				
	z	m/z (calculated)	m/z (observed)	ppm error
	3	1148.1635	1148.1562	-6.3
N + 1	4	861.6233	861.6223	-1.16
	5	688.6959	688.6979	2.97
	6	**573.5781**	**573.5782**	0.31
	8	471.0633	471.0636	0.57
	6	628.4202	628.4203	0.18
N +2	5	754.3057	754.3070	1.79
	4	943.3848	943.3850	0.20
	3	1258.1821	1258.1803	-1.42

3.3.3. Effect of lipids on the cNMP-based primer extension reactions under DH-RH conditions

Encouraged by the addition of two intact nucleotides in the aforementioned primer extension reactions involving 2′, 3′ cNMP, we investigated the effect of lipids on these reactions. Previously, the presence of lipids has been demonstrated to protect the molecules involved in the nonenzymatic oligomerization and template-directed primer extension reactions, against hydrolysis. Further, their presence also imparts an identity to the encapsulated system that form instantaneously post rehydration in such scenarios [20,26,28]. Given this, primer extension reactions were also performed in the presence of 5 mM POPC for the template-directed as well as untemplated reactions, using both OH-primer and NH_2-primer.

As shown in Figure 3.3 (panels a-d), extension till two bases ('N+1' and 'N+2') was observed in all the reactions. The presence of lipids was observed to enhance the yield of extended products in both the templated as well as untemplated reactions (Figure 3.3e and 3.3f). In the case of POPC-assisted template-directed OH-primer

reactions, yield of the extended product enhanced from 12.9% (in the absence of POPC) to 18.4 % within one DH-RH cycle, and 15.5% of the product persisted even after five DH-RH cycles (Figure 3.3e and 3.3b). Similarly, in the untemplated reactions, the presence of POPC led to an increase in the product yield to 13.3% (as compared to 8% in the absence of POC) within one DH-RH cycle, wherein 9.6% of it remained even after five DH-RH cycles. In the NH_2-primer containing POPC-assisted reactions, the yield enhanced to 13.6% (12.4% in the absence of POPC) and 21.6% (15.9% in the absence of POPC) in untemplated and template-directed reactions, respectively (Figure 3.3f and 3.3a). Up to 9.5% (in untemplated) and 19.8% (in template-directed) of the product persisted even after five DH-RH cycles in the presence of POPC, whereas only 7.7 % and 11.9 % was observed in its absence (in untemplated and template-directed reactions, respectively). This increase in the extended primer yield in the presence of POPC could be because of its protecting effect towards the substrates (primer as well as cyclic nucleotides) against hydrolysis. This, in turn, would have enhanced the availability of the reacting substrates, thereby leading to an enhancement in the product yield.

Importantly, as mentioned earlier, we investigated the effect of prebiotically realistic hybrid compartments on our enzyme-free template-directed primer extension reactions. Towards this, prebiotically relevant composite compartments comprising POPC along with equimolar concentrations of either oleic acid (OA) or glycerol monooleate (GMO), were investigated (see Methods for details). Similar to the POPC-assisted reactions, both POPC:OA and POPC:GMO systems were observed to enhance the yield of the extended product when compared to the reactions without any lipid in them (Figure 3.4, Table 3.2). Nonetheless, the difference between the yield of extended products in the presence of any of the membrane systems was found to be insignificant (based on a two-tailed t-test). These results signify that all the lipid compositions evaluated seemed to impart protection to the extended products against hydrolysis as the amount of extended product did not decrease significantly (based on a two-tailed t-test) even after five DH-RH cycles

Template	-				+					-				+			
Primer	OH									NH₂							
cNMP	2′-3′ cAMP									2′-3′ cAMP							
Time pt.	0	1	2	5	0	1	2	5		0	1	2	5	5	2	1	0
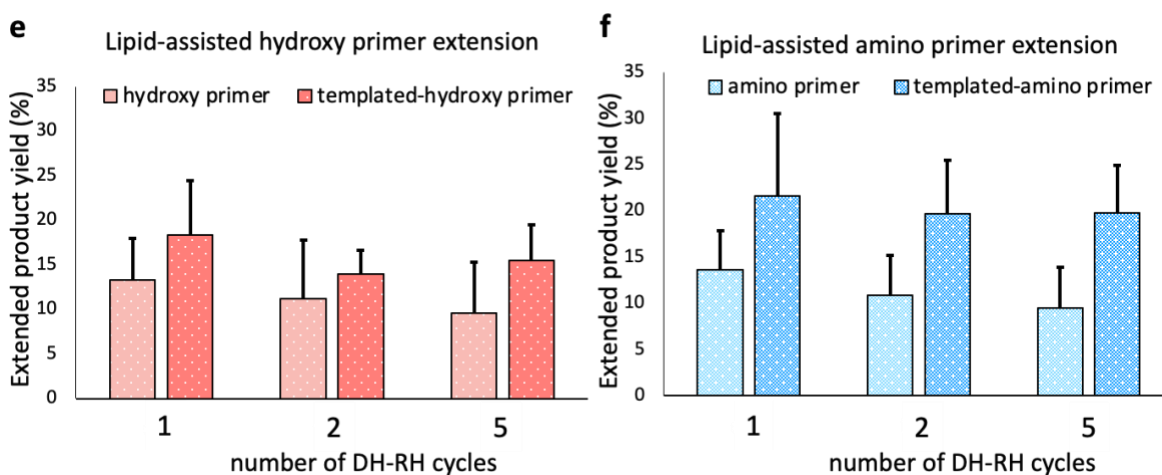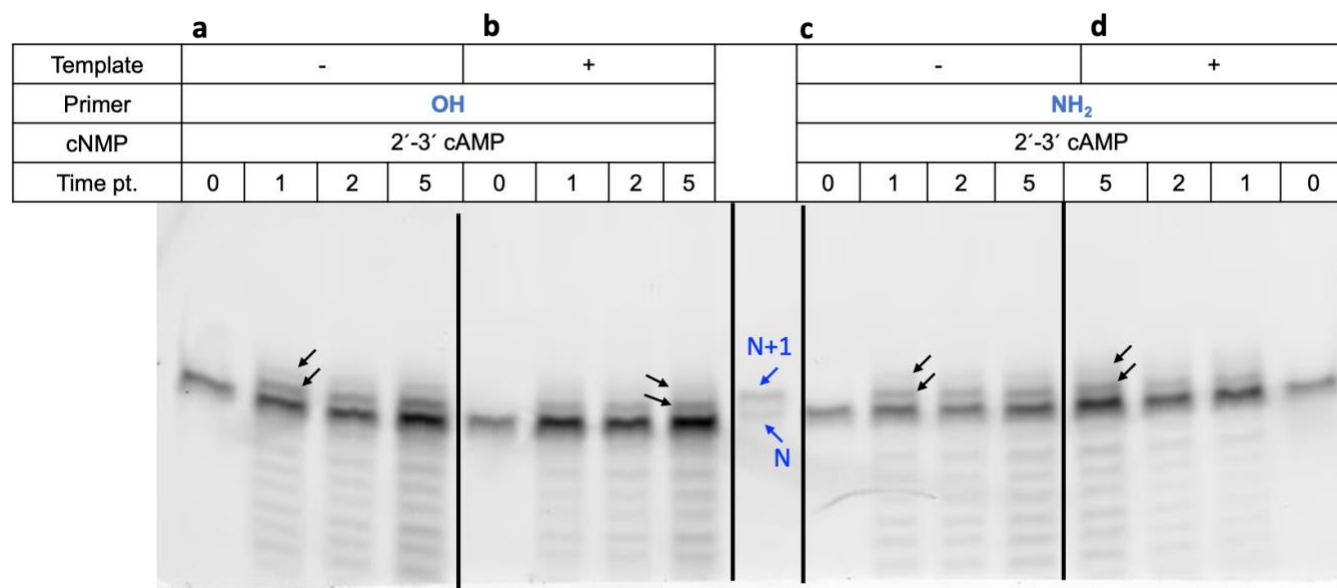
Figure 3.3: Effect of POPC on RNA primer extension. Reactions were performed using template U, 2′, 3′ cAMP, and either NH₂-primer or OH-primer, over repeated cycles of DH-RH (i.e., cycle 0, cycle 1, cycle 2 and cycle 5). The black arrows indicate the extended product. In the N+1 lane, 'N' indicates the control 20-mer RNA primer while 'N + 1' indicates extension of the primer by one nucleotide. The black vertical lines in the above two panels have been used to demarcate two reaction sets that were run on the same gel. Panels e and f show the quantified total yield (%) of the extended primer in the lipid-assisted reactions of OH-primer (panel e) and NH₂-primer (panel f), respectively. Yields were quantified for both the untemplated and

70

template-directed reactions (as depicted in the legend), after varying the number of DH-RH cycles (x-axis). The difference was found to be insignificant based on a two-tailed t-test. Error bars = s.d., N=3.
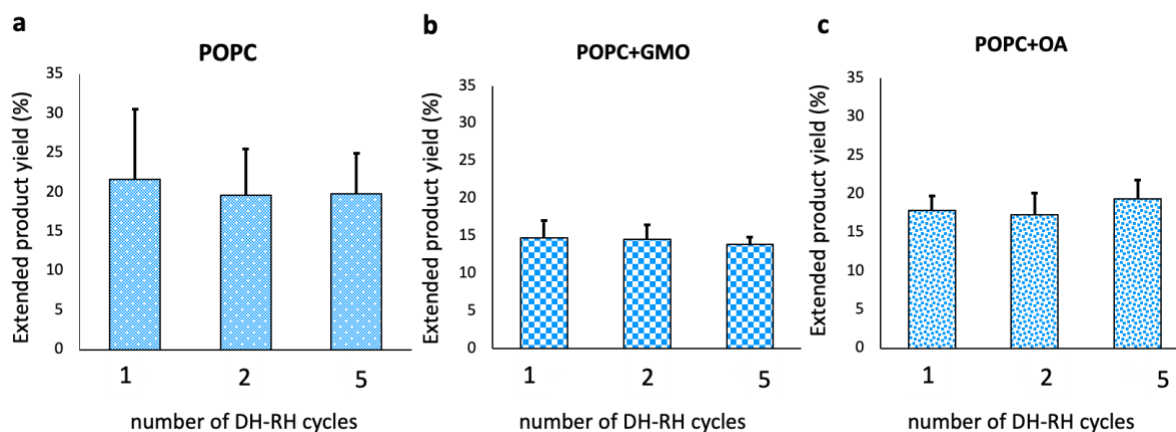
Figure 3.4: Effect of different lipids on lipid-assisted RNA primer extension.
Reactions were performed using template U, 2′, 3′ cAMP and NH_2-primer, in the presence of 5 mM lipid, over repeated cycles of DH-RH, as indicated by X-axis. Different panels show the quantified yield (%) of extended product in the presence of compartments with different lipid compositions, viz. pure POPC (a), 1:1 POPC:GMO (b) and 1:1 POPC:OA (c). The difference was found to be insignificant based on a two-tailed t-test. Error bars = s.d., N=3.

Table 3.2: The yield (%) of the extended product in the presence of different lipid systems i.e., pure POPC and two hybrid membrane systems (1:1 POPC:GMO and 1:1 POPC:OA) over repeated DH-RH cycles

Lipid-assisted template-directed primer extension using cAMP			
	number of DH-RH cycles		
	1	2	5
POPC	21.61 ± 8.91	19.67 ± 5.83	19.80 ± 5.14
POPC + GMO	14.73 ± 2.32	14.58 ± 1.92	13.92 ± 0.92
POPC + OA	17.83 ± 1.93	17.31 ± 2.85	19.38 ± 2.48

3.3.4. Primer extension using a pyrimidine-based cyclic nucleotides (cCMP) under DH-RH conditions

Our results showed template-directed primer extension using cAMP as the incoming nucleotide, against uracil as the templating base. Importantly, purines are known to stack better and potentially have a better tendency to undergo concentration-dependent condensation reactions. Similar results were also observed in the nonenzymatic oligomerizations reactions, where the yield of oligomers in cCMP reactions were lower than that of cAMP [28]. Therefore, we examined how the presence of a pyrimidine-based cNMP i.e., cCMP, would influence these reactions as compared to purine-based cAMP. Towards this, we performed template-directed primer extensions using cyclic 2′, 3′ cCMP (Figure 3.5) against template G. The reactions were performed with both OH-primer and NH$_2$-primer.

As observed in Figure 3.5a, the primer extension with up to two nucleotide additions was observed for both of these primers when analyzed using PAGE (Figure 3.5a). The intact nucleotide additions in the primer extension reactions (involving a 10-mer primer) were further confirmed using LC-MS. The m/z corresponding to 'N+1' and 'N+2' was observed within less than 5 ppm error (Table 3.3). Upon quantification, 3.5% and 3.4% of the extended product was observed within one DH-RH cycle in the case of OH-primer and NH$_2$-primer, respectively (Figure 3.5a, 3.5c and 3.5d). With an increase in the number of DH-RH cycles to five, the yield of the extended product decreased significantly (based on a two-tailed t-test) for both the OH-primer (1.32%) and NH$_2$-primer (1.9%). The reason behind this could be the instability of RNA under our reaction conditions i.e., 90°C and pH 8 for five DH-RH cycles.

As alluded to earlier, the presence of lipids under such scenarios has been shown to impart protection against hydrolysis. In order to investigate the effect of lipids on template-directed primer extension, the extension reactions were performed using cCMP against template G in the presence of POPC. As observed in Figure 3.5b and table 3.4, two extension bands were seen in the lipid-assisted extension of both OH-primer and NH$_2$-primer within one DH-RH cycle. Both the extension bands persisted

till five DH-RH cycles in these reactions. Upon quantification, the yield of extended product in these POPC-assisted reactions was observed to be significantly higher (based on a two-tailed t-test) than in the reactions without lipid. In the case of POPC-assisted OH-primer reactions, 7.0 % of the extended product was observed within one DH-RH cycle (Figure 3.5e) as compared to just 3.5 % in the corresponding reactions without lipid. Moreover, the decrease in the extended product yield after five DH-RH cycles was found to be insignificant (based on a two-tailed t-test) in POPC-assisted OH-primer reactions. Conversely, in the reactions without lipids, with the increasing number of DH-RH cycles, the yield of the extended product significantly decreased.
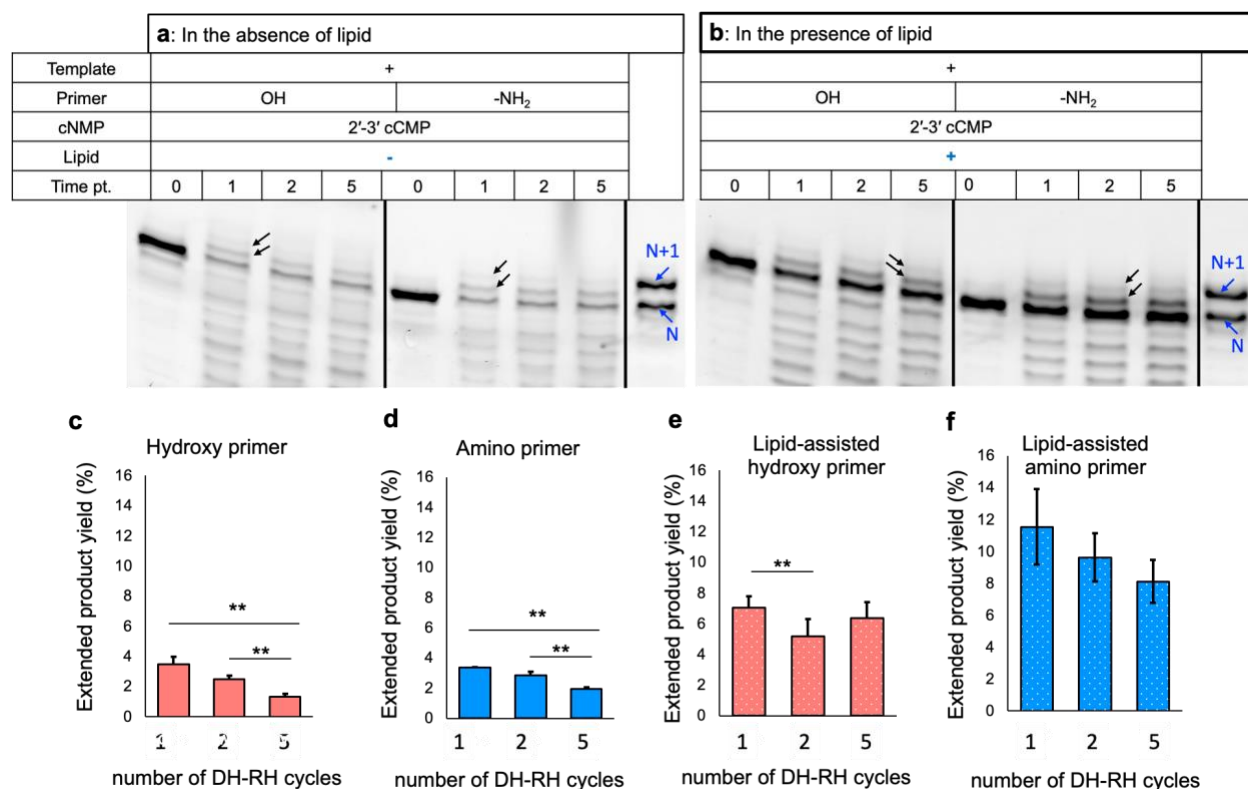

Figure 3.5: Primer extension using cCMP under multiple DH-RH cycles and effect of lipids on these reactions. Reactions were performed using template G, 2′, 3′ cCMP and with NH_2-primer or OH-primer, over repeated DH-RH cycles in the absence (a) or in the presence of 5 mM POPC (b). The black arrows indicate the

extended products. In the N+1 lane, 'N' indicates the control 20-mer RNA primer while 'N + 1' indicates extension of the primer by one nucleotide. The black vertical lines in the above two panels have been used to demarcate two reaction sets that were run on the same gel. N=3. Panels c to f show the quantified total yield (%) of the extended primer in the absence of POPC (c and d) and POPC-assisted reactions (e and f), for OH-primer and NH_2-primer, respectively. Yields were quantified for both untemplated and templated reactions (as depicted in the legend) after varying number of DH-RH cycles (x-axis). The ** indicates a significant change with a p-value<0.01 based on a two-tailed t-test. The difference was found to be insignificant for the remaining bars. Error bars = s.d., N=3.

Table 3.3: Masses observed during LC-MS characterization of primer extension reactions using 2′, 3′ cCMP, along with calculated parts per million (ppm) error. Masses for one nucleotide and two nucleotide extensions on to the 10-mer OH-primer have been indicated as 'N+1' and 'N+2', respectively.

2′-3′ cCMP				
	z	m/z (calculated)	m/z (observed)	ppm error
N +1	2	1710.7433	1710.73666	-3.89
	3	1140.1598	1140.15784	-1.7
	4	854.8680	854.8696	1.87
	7	488.064321	488.064456	0.28
N +2	1	3728.5385	3728.5309	-2.03
	2	1863.7656	1863.7699	2.32
	3	1242.1746	1242.1729	-1.39
	5	744.9019	744.9059	5.34
	7	531.7850	531.7852	0.36
	4	931.3792	931.3779	-1.36
	6	620.5837	620.5846	1.39

Table 3.4: The yield (%) of the extended product in the presence and absence of POPC using 2′, 3′ cCMP and with NH₂-primer or OH-primer over repeated DH-RH cycles

Template-directed primer extension using cCMP			
	number of DH-RH cycles		
	1	2	5
OH-primer	3.47 ± 0.5	2.49 ± 0.25	1.32 ± 0.2
NH₂-primer	3.36 ± 0.04	2.85 ± 0.22	1.95 ± 0.09
POPC-assisted OH-primer	7.04 ± 0.74	5.19 ± 1.11	6.37 ± 1.03
POPC-assisted NH₂-primer	11.55 ± 2.36	9.63 ± 1.52	8.12 ± 1.35

In the case of POPC-assisted NH₂-primer reactions, 11.6 % of extended primer was observed within one DH-RH cycle for NH₂-primer (Figure 3.5f). This was significantly higher when compared to the 3.4 % yield obtained in the absence of POPC (Figure 3.5d). Moreover, the decrease in the extended product yield after five DH-RH cycles was found to be insignificant (based on a two-tailed t-test) in the POPC-assisted NH₂-primer reactions (Figure 3.5f). These results emphasize the fact that the yield and stability (against hydrolysis) of the extended primer was enhanced significantly based on a two-tailed t-test in the lipid-assisted reactions.

3.4. Discussion

Enzyme-free oligomerization of RNA and its propagation, are central to the RNA World Hypothesis [2]. Previous studies in this context have predominantly employed activated nucleotides to study these phenomena [4,41–47]. However, the availability of activated nucleotides in significant amounts on the prebiotic Earth is debatable [28,48]. In this context, a related study demonstrated template-directed primer extension reactions using non-activated 5′-NMPs under acidic terrestrial geothermal conditions i.e., acidic pH, high temperature and DH-RH cycles [20]. However, systematic characterization of these extended products indicated abasic sites due to the cleavage of the glycosidic bond especially because of the acidic pH [26,27]. Since cNMPs are intrinsically reactive due to the ring strain present in them, they do not

require harsh conditions such as acidic pH for undergoing oligomerization reactions [49]. In an earlier study, we demonstrated the nonenzymatic oligomerization of these cNMPs under high temperature, alkaline conditions (pH 8), using repetitive dry-wet cycles [28]. In this study, we demonstrate the enzyme-free template-directed primer extension using cNMPs, under the same conditions. Our results show primer extension happening, with up to two nucleotide additions in all the reactions investigated, using 2′, 3′cNMP under alkaline conditions (pH 8 and 10). Nonetheless, alkaline conditions are known to be challenging for RNA stability as it leads to phosphodiester bond cleavage. This explains the optimal yield of the extended product at pH 8. However, similar reactions with 3′, 5′ cNMP did not yield in any extension products under our experimental conditions. This could be due to their relatively higher stability towards hydrolysis when compared to 2′, 3′ cAMP [28,49]. Also, 3′, 5′ cNMP contains a six-membered ring when compared to a five-membered ring in 2′, 3′ cNMP. This makes it comparatively less reactive due to the lower ring strain that it experiences (8.9 kcal/mol) when compared to the five-membered ring of 2′,3′ cAMP (12 kcal/mol, [49].

Pertinently, the extension of primer was observed with both OH and NH_2-primers, for both cAMP and cCMP containing reactions. This signifies the generalizability of primer extension using cNMPs without the imminent need for any external activation for facilitating these reactions. Upon comparing the yield of the extended product, the cAMP containing template-directed reactions were observed to be significantly higher yielding as compared to the corresponding cCMP reactions (Table 3.5). In the case of template-directed reactions using OH-primer, the difference in cAMP and cCMP reactions, after one and two DH-RH cycles, was found to be insignificant based on a two-tailed t-test (Table 3.5). After five DH-RH cycles, the extended product in cAMP was found to be significantly (Table 3.5) higher. In template-directed NH_2-primer extension reactions, after one and two DH-RH cycles, the extended product was significantly higher in cAMP (Table 3.5). However, by five DH-RH cycles, the difference was found to be insignificant (Table 3.5). In lipid-assisted reactions, the primer extension was observed to occur in the presence of POPC as well as

prebiotically relevant hybrid membrane systems (POPC+GMO and POPC+OA), highlighting the compatibility of these reactions with model protocellular membranes. In the case of cCMP-based reactions, the presence of PL was observed to significantly increase the yield of the extended product. This could be because of the capability of lipid to impart protection to RNA as well as cNMPs against hydrolysis [28].

Table 3.5: Comparison of template-directed RNA primer extension using 2′, 3′ cAMP vs 2′, 3′ cCMP in either without lipid or lipid-assisted reactions, across various DH-RH cycles between the two reactions, using two-tailed type 2 t-test. (Significant p-value is highlighted with a green background). N = 3.

Cycles being compared	2′, 3′ cAMP vs 2′, 3′ cCMP							
	Template -directed							
	Without lipid				Lipid-assited			
	NH$_2$-primer		OH-primer		NH$_2$-primer		OH-primer	
	p-value		p-value		p-value		p-value	
1 with 1	0.04	*	0.06	ns	0.077	ns	0.13	*
2 with 2	0.037	*	0.09	ns	0.02	*	0.003	**
5 with 5	0.09	ns	0.016	*	0.009	**	0.008	**

Most previous studies have employed activated nucleotides to study primer extension reactions, and these are facilitated at ambient temperatures [6,41,46,50–52]. Therefore, one of the crucial challenges faced by enzyme-free template-directed primer extension reactions under these ambient conditions, is "strand separation" [42]. This is essential for the replicated duplex strands to separate, in order to allow for multiple rounds of information propagation. Towards this, few studies that involved the use of activated nucleotides, have suggested workarounds including the use of viscous solvent, pH change or incorporating backbone heterogeneity in order to overcome this problem [42,53,54]. In our study, the template-directed primer extension reactions occur under prebiotically relevant conditions of high temperature (90°C), and under

geochemically pertinent DH-RH cycles. Such a high temperature (90°C) is expected to separate the template-primer complex; however, previous computational studies have argued that under dry conditions, the melting temperature of a duplex with the length comparable to our template-primer combination (i.e., 9 mismatch in a 30 base-paired duplex) is higher than 90°C[55–57]. This argument is further supported by the fact that we observed 12.9% and 8% of the extended product in the template-directed and untemplated reactions, respectively, which essentially translates to a 61.25% increase within one DH-RH cycle in the case of reactions containing OH-primer and 2'-3' cAMP. Thus, we hypothesise that recurrent DH-RH fluctuations at 90°C can, in principle, facilitate the annealing of primer and template to form duplex during the "dehydrated" phase and assist in strand separation in the "rehydrated" phase due to high temperature aqueous conditions. Additionally, 2′, 3′ cNMP involve intramolecular phosphodiester bonds, which upon oligomerization and primer extension, result in extended products with inherently mixed intermolecular phosphodiester linkages. Such random backbone heterogeneity in the duplexes is known to result in their melting at lower temperatures [53], further facilitating in effective strand separation. Given these direct implications for RNA propagation on the early Earth, our study underlines the importance of facilitating template-directed primer extension reactions, using intrinsically active cNMPs under prebiotically pertinent terrestrial geothermal conditions. Further, our results also highlight the intrinsic capability of environmental fluctuations such as DH-RH cycles, in catalyzing nonenzymatic information propagation on the early Earth, thus renewing our understanding about how oligomerization and propagation of RNA would have occurred under realistic prebiotic conditions, during the emergence of a putative RNA World.

3.5. References

1.   Stano, P. & Mavelli, F. Protocells Models in Origin of Life and Synthetic Biology. *Life* **5**, 1700–1702 (2015).

2.   Gilbert, W. Origin of life: The RNA world. *Nature* **319**, (1986).

3.   Li, L. *et al.* Enhanced Nonenzymatic RNA Copying with 2-Aminoimidazole

Activated Nucleotides. *J. Am. Chem. Soc.* **139**, 1810–1813 (2017).

4.   Walton, T., Zhang, W., Li, L., Tam, C. P. & Szostak, J. W. The Mechanism of Nonenzymatic Template Copying with Imidazole-Activated Nucleotides. *Angew. Chemie Int. Ed.* **58**, 10812–10819 (2019).

5.   Zhou, L. *et al.* Non-enzymatic primer extension with strand displacement. *Elife* **8**, (2019).

6.   Deck, C., Jauker, M. & Richert, C. Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA. *Nat. Chem.* **3**, 603–608 (2011).

7.   Kaiser, A. & Richert, C. Nucleotide-Based Copying of Nucleic Acid Sequences without Enzymes. *J. Org. Chem.* **78**, 793–799 (2013).

8.   Lohrmann, R. & Orgel, L. E. Prebiotic Activation Processes. *Nature* **244**, 418–419 (1973).

9.   Blain, J. C. & Szostak, J. W. Progress Toward Synthetic Cells. *Annu. Rev. Biochem.* **83**, 615–640 (2014).

10.   Szostak, J. W., Bartel, D. P. & Luisi, P. L. Synthesizing life. *Nature* **409**, 387–390 (2001).

11.   Orgel, L. E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **39**, 99–123 (2004).

12.   Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239–242 (2009).

13.   Powner, M. W., Sutherland, J. D., Szostak, J. W. & Matthew W. Powner,*,† John D. Sutherland, ‡ and Jack W. Szostak†. Chemoselective multicomponent one-pot assembly of purine precursors in water. *J. Am. Chem. Soc.* **132**, 16677–16688 (2010).

14.   Sidney Becker,* Ines Thoma,* Amrei Deutsch, Tim Gehrke, Peter Mayer, Hendrik Zipse, T. C. A high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. **352**, (2016).

15.   Lehninger, A., Nelson, D. L. & Cox, M. M. *Lehninger-Principles of Biochemistry-Fourth Edition. W. H. Freeman and Company* (2008). doi:10.1146/annurev-arplant-050718-100353

16.   Yi, R., Hongo, Y. & Fahrenbach, A. C. Synthesis of imidazole-activated

ribonucleotides using cyanogen chloride. *Chem. Commun.* 1–4 (2018). doi:10.1039/C7CC08489G

17.   Sosson, M. & Richert, C. Enzyme-free genetic copying of DNA and RNA sequences. *Beilstein J. Org. Chem.* **14**, 603–617 (2018).

18.   Jauker, M., Griesser, H. & Richert, C. Copying of RNA Sequences without Pre-Activation. *Angew. Chemie - Int. Ed.* **54**, 14559–14563 (2015).

19.   Röthlingshöfer, M. & Richert, C. Chemical primer extension at submillimolar concentration of deoxynucleotides. *J. Org. Chem.* **75**, 3945–3952 (2010).

20.   Bapat, N. V. & Rajamani, S. Templated replication (or lack thereof) under prebiotically pertinent conditions. *Sci. Rep.* **8**, 15032 (2018).

21.   Higgs, P. G. The effect of limited diffusion and wet–dry cycling on reversible polymerization reactions: Implications for prebiotic synthesis of nucleic acids. *Life* **6**, 1–16 (2016).

22.   Ross, D. & Deamer, D. Dry/Wet Cycling and the Thermodynamics and Kinetics of Prebiotic Polymer Synthesis. *Life* **6**, 28 (2016).

23.   Damer, B. & Deamer, D. Coupled Phases and Combinatorial Selection in Fluctuating Hydrothermal Pools: A Scenario to Guide Experimental Approaches to the Origin of Cellular Life. *Life* **5**, 872–887 (2015).

24.   Toppozini, L., Dies, H., Deamer, D. W. & Rheinstädter, M. C. Adenosine Monophosphate Forms Ordered Arrays in Multilamellar Lipid Matrices: Insights into Assembly of Nucleic Acid for Primitive Life. *PLoS One* **8**, 62810 (2013).

25.   Sarkar, S., Dagar, S. & Rajamani, S. Influence of Wet–Dry Cycling on the Self-Assembly and Physicochemical Properties of Model Protocellular Membrane Systems. *ChemSystemsChem* **3**, (2021).

26.   Mungi, C. V., Rajamani, S., Mungi, C. V. & Rajamani, S. Characterization of RNA-Like Oligomers from Lipid-Assisted Nonenzymatic Synthesis: Implications for Origin of Informational Molecules on Early Earth. *Life (Basel, Switzerland)* **5**, 65–84 (2015).

27.   An, R. *et al.* Non-Enzymatic Depurination of Nucleic Acids: Factors and Mechanisms. *PLoS One* **9**, e115950 (2014).

28.   Dagar, S., Sarkar, S. S. & Rajamani, S. Geochemical influences on nonenzymatic

oligomerization of prebiotically relevant cyclic nucleotides. *RNA* **26**, 756–769 (2020).

29.   Tapiero, C. M. & Nagyvary, J. Prebiotic Formation of Cytidine N ucleotides. *Nature* **231**, 42–43 (1971).

30.   Verlander, M. S., Lohrmann, R. & Orgel, L. E. Catalysts for the self-polymerization of adenosine cyclic 2',3'-phosphate. *J. Mol. Evol.* **2**, 303–316 (1973).

31.   Verlander, M. S. & Orgel, L. E. Analysis of High Molecular-Weight Material From Polymerization of Adenosine Cyclic 2',3'-Phosphate. *J. Mol. Evol.* **3**, 115–120 (1974).

32.   Costanzo, G., Saladino, R., Crestini, C., Ciciriello, F. & Di Mauro, E. Nucleoside phosphorylation by phosphate minerals. *J. Biol. Chem.* **282**, 16729–16735 (2007).

33.   Suárez-Marina, I. *et al.* Integrated synthesis of nucleotide and nucleosides influenced by amino acids. *Commun. Chem.* **2**, 1–8 (2019).

34.   Saladino, R., Botta, G., Pino, S., Costanzo, G. & Di Mauro, E. From the one-carbon amide formamide to RNA all the steps are prebiotically possible. *Biochimie* **94**, 1451–1456 (2012).

35.   Rajamani, S. *et al.* Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Orig. Life Evol. Biosph.* **38**, 57–74 (2008).

36.   Bapat, N. V. Role of co-solutes in nonenzymatic RNA replication on prebiotic Earth Indian Institute of Science Education and. (2018).

37.   Mansy, S. S. Membrane transport in primitive cells. *Cold Spring Harb. Perspect. Biol.* **2**, 1–15 (2010).

38.   Jin, L., Kamat, N. P., Jena, S. & Szostak, J. W. Fatty Acid/Phospholipid Blended Membranes: A Potential Intermediate State in Protocellular Evolution. *Small* **1704077**, 1–9 (2018).

39.   Costanzo, G. *et al.* Non-Enzymatic Oligomerization of 3', 5' Cyclic AMP. *PLoS One* **11**, (2016).

40.   Wunnava, S. *et al.* Acid-catalyzed RNA-oligomerization from 3´,5´-cGMP. *Chem. – A Eur. J.* (2021). doi:10.1002/chem.202103672

41.   Adamala, K. & Szostak, J. W. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science (80-.).* **342**, 1098–1100 (2013).

42. Mariani, A., Bonfio, C., Johnson, C. M. & Sutherland, J. D. pH-Driven RNA Strand Separation under Prebiotically Plausible Conditions. *Biochemistry* **57**, 6382–6386 (2018).

43. O'Flaherty, D. K., Zhou, L. & Szostak, J. W. Nonenzymatic Template-Directed Synthesis of Mixed-Sequence 3′-NP-DNA up to 25 Nucleotides Long Inside Model Protocells. *J. Am. Chem. Soc.* **141**, 10481–10488 (2019).

44. Zhang, S. J., Duzdevich, D. & Szostak, J. W. Potentially Prebiotic Activation Chemistry Compatible with Nonenzymatic RNA Copying. *J. Am. Chem. Soc.* **142**, 14810–14813 (2020).

45. Joyce, G. F. & Szostak, J. W. Protocells and RNA Self-Replication. *Cold Spring Harb. Perspect. Biol.* **10**, a034801 (2018).

46. Walton, T., Zhang, W., Li, L., Tam, C. P. & Szostak, J. W. The Mechanism of Nonenzymatic Template Copying with Imidazole-Activated Nucleotides. *Angew. Chemie Int. Ed.* **58**, 10812–10819 (2019).

47. Prywes, N., Blain, J. C., Del Frate, F. & Szostak, J. W. Nonenzymatic copying of RNA templates containing all four letters is catalyzed by activated oligonucleotides. *Elife* **5**, 1–14 (2016).

48. Olasagasti, F. & Rajamani, S. Lipid-assisted polymerization of nucleotides. *Life* **9**, 1–10 (2019).

49. Khorana, H. G., Tener, G. M., Wright, R. S. & Moffatt, J. G. Cyclic Phosphates. III. Some General Observations on the Formation and Properties of Five-, Six- and Seven-membered Cyclic Phosphate Esters. *J. Am. Chem. Soc.* **79**, 430–436 (1957).

50. Motsch, S., Pfeffer, D. & Richert, C. 2′/3′ Regioselectivity of Enzyme-Free Copying of RNA Detected by NMR. *ChemBioChem* **21**, 2013–2018 (2020).

51. Bapat, N. V. & Rajamani, S. Effect of Co-solutes on Template-Directed Nonenzymatic Replication of Nucleic Acids. *J. Mol. Evol.* **81**, 72–80 (2015).

52. Duzdevich, D., Carr, C. E. & Szostak, J. W. Deep sequencing of non-enzymatic RNA primer extension. *Nucleic Acids Res.* **48**, e70–e70 (2020).

53. Engelhart, A. E., Powner, M. W. & Szostak, J. W. Functional RNAs exhibit tolerance for non-heritable 2′–5′ versus 3′–5′ backbone heterogeneity. *Nat. Chem.*

5, 390–394 (2013).

54. He, C., Lozoya-Colinas, A., Gállego, I., Grover, M. A. & Hud, N. V. Solvent viscosity facilitates replication and ribozyme catalysis from an RNA duplex in a model prebiotic process. *Nucleic Acids Res.* **47**, 6569–6577 (2019).

55. Roy, S., Bapat, N. V., Derr, J., Rajamani, S. & Sengupta, S. Emergence of ribozyme and tRNA-like structures from mineral-rich muddy pools on prebiotic earth. *J. Theor. Biol.* **506**, 110446 (2020).

56. Dumousseau, M., Rodriguez, N., Juty, N. & Novère, N. Le. MELTING, a flexible platform to predict the melting temperatures of nucleic acids. *BMC Bioinformatics* **13**, 101 (2012).

57. Le Novere, N. MELTING, computing the melting temperature of nucleic acid duplex. *Bioinformatics* **17**, 1226–1227 (2001).

Chapter 4

Nonenzymatic oligomerization using adenosine 5′-triphosphate under wet-dry cycles

4. 1 Introduction

The well-explored "RNA World Hypothesis" suggests RNA be the first biomolecule to have emerged on the early Earth[1]. This would have aided the transition from a "chemical" world to a "biological" one. The concept of RNA World presumes the presence of long strands of RNA, which would have served as genetic information as well as a catalyst on the prebiotic Earth. Extant biology makes use of highly specific and efficient proteins like polymerases that utilize 5′-nucleoside triphosphates (NTPs) to facilitate phosphodiester bond formation to form RNA[2]. However, on early Earth, such complex protein machinery would have been implausible. Thus, the emergence of long strands of RNA is thought to have occurred by nonenzymatic means and driven by environmental constraints[3–5]. Earlier studies have predominantly employed activated nucleotides such as 5' phosphorimidazole-activated nucleotides (ImpNs) to demonstrate enzyme-free oligomerization of nucleotides[6–9]. Although insightful, the presence of ImpNs in significant amounts on the prebiotic Earth is debatable[10]. Given this, nonenzymatic oligomerization of non-activated nucleotides such as 5′-nucleoside monophosphates (5′-NMPs) was attempted under terrestrial acidic geothermal conditions[3,4]. The characterization of the resultant oligomers showed the formation of oligomers with abasic sites due to the cleavage of the glycosidic bond[3]. Other relevant studies demonstrated the oligomerization of cyclic nucleotides (cNMPs) under terrestrial hydrothermal conditions[5,11–14].

The characteristic feature of such aforementioned terrestrial sites is the environmental fluctuations that it supports including temperature and seasonal changes. These fluctuations give rise to recurrent dry-wet or dehydration/rehydration (DH-RH) cycles[15]. Dry phase concentrates the solute molecules present in such scenarios, thus, overcoming the problem of dilution. The wet phase redistributes the molecules and also provides a medium for collisions to occur. Therefore, such scenarios provide a kinetic trap and are shown to facilitate concentration-dependent reactions. However, such

enzyme-free oligomerization studies utilizing NTPs are lacking. The availability of NTPs on the early Earth was controversial due to limitations in the availability of soluble phosphate, known as the "phosphate-problem"[16,17]. In this regard, few previous studies have demonstrated the phosphorylation of nucleosides to yield polyphosphates on the surface of phosphide minerals, overcoming the requirement of soluble phosphorus[18–20]. Another study has demonstrated the presence of soluble phosphorus in significant amounts in a carbonate-rich lake[21]. Recently, the formation of NTPs in the presence of nickel borate under prebiotically relevant conditions was also shown[22]. All these studies emphasize the plausibility of NTPs to result in the formation of RNA on the early Earth. Subsequently, another relevant study demonstrated a one-pot generation and utilization of NTPs by proteinaceous and ribozyme polymerases, for template-directed primer extensions[23]. The recognition and utilization of NTPs is a fundamental feature of extant life too, thus making this study an exciting one as it demonstrates the formation of RNA from NTPs in an "abiotic" world too. Metal ions in particular, are known for their phosphoryl transfer activity[24,25]. For e.g., extant proteinaceous polymerases employ magnesium ions as cofactors in order to catalyze phosphodiester formation[2,26]. Magnesium ions facilitates phosphodiester bond formation by decreasing the repulsion between negatively charged nucleotides and enhancing the nucleophilicity of the incoming ribose hydroxyl group.

In this backdrop, we aimed to investigate the effect of the presence of different prebiotically relevant ions, on the nonenzymatic oligomerization of 5′-adenosine triphosphate (ATP). Initially, the effect of experimental conditions such as temperature and pH were evaluated on the oligomerization of ATP under DH-RH cycling conditions. A high temperature of 90°C and pH 7 gave an optimum yield of intact oligomers up to trimers. Subsequently, we investigated the effect of prebiotically relevant metal ions including Mg^{2+}, Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}, on the nonenzymatic oligomerization reactions of ATP. To begin with, sulfate salts of all the metal ions were used. Magnesium sulfate resulted in the highest yield of intact oligomers, which were stable till five DH-RH cycles. Following this, the effect of the anion was also evaluated by using magnesium chloride instead of magnesium sulfate. Magnesium chloride was

observed to be better at facilitating the oligomerization reaction. Altogether, this study provides fundamental insights towards the formation of RNA under early Earth conditions using NTPs as the monomeric moieties. This also sheds light on the potential link between abiotic synthesis of NTP and the biotic synthesis of RNA by ribozyme polymerases, both of which cannot occur effectively without the presence of magnesium ions.

4.2. Experimental Section

4.2.1 Materials

100 mM solution of the disodium salt of 5′-adenosine triphosphate (ATP) in 10 mM Tris buffer at pH 7, was purchased from Sigma–Aldrich and used without further purification. All other reagents used were of analytical grade and purchased from Sigma–Aldrich.

4.2.2 Methods

4.2.2.1. Simulating early Earth conditions

Early Earth conditions were simulated using a bench-top heating block that was maintained at high temperature, that is at 90°C. The oligomerization reactions were carried out in 20 ml glass vials purchased from Chemglass whose caps are fitted with PTFE septa. The anoxygenic environment was maintained by delivering a gentle flow of nitrogen into the vials through two PEEK tubings of about 1–1.5 in, one acting as an inlet and another as an outlet for the gas.

4.2.2.2. Experimental Design

A typical reaction mixture consisted of 5 mM ATP (stock = 100 mM stock in 10 mM Tris pH 7) in nanopure water, with the final pH set at ~7.

Aqueous heating control reactions: The reaction mixture was sealed by an air-tight cap and incubated at 90°C. The samples were collected after various time

periods i.e., immediately at the initiation of the reactions (0 hr.), after 1, 4, 8 and 12 hrs. and analyzed.

Dry heating control reactions: The reaction mixture was distributed in different vials for each time point (i.e., 0 hr., 1, 4, 8 and 12 hrs) and dried at 90°C under a gentle flow of N_2. After each time period, one vial was collected and the dried reaction mixture was rehydrated with nanopure water and analyzed.

Nonenzymatic oligomerization of ATP under DH-RH conditions: The reaction mixture was subjected to multiple DH-RH cycles by drying it at 90°C under a gentle flow of N_2 followed by rehydration with 1 mM ATP solution, to compensate for the hydrolysis of the substrate. The contents were mixed properly followed by a prolonged dehydration phase during each DH–RH cycle, with 1h per cycle. The rehydrated samples were collected at different time periods including, cycle 0, cycle 1, cycle 3 and cycle 5, and analyzed using HPLC and LC-MS/MS.

4.2.2.3. High Performance Liquid Chromatography (HPLC) analysis

HPLC analysis was performed using the Infinity series 1260 HPLC instrument from Agilent Technologies, Santa Clara, CA, USA. After removing the lipids using previously standardized butanol-hexane extraction method, the reaction products in the aqueous phase were loaded on to the HPLC [3]. The molecules were separated by an anion-exchange column viz. DNAPac PA 200 (Thermo Scientific, Sunnyvale, CA, USA). It separates molecules based on their interaction with the column through the phosphate moiety, thus providing single-nucleotide resolution. The separation method was standardized and the samples were run using an elution system comprising of 2 mM Tris buffer at pH 8 (Solvent A) and 0.4 M sodium perchlorate in 2 mM Tris buffer at pH 8, (Solvent B). The gradient used was as follows: 0% solvent B for 3 minutes, from 3 min to 10 min solvent B goes from 0 to 30%, in the next 3 min (i.e., 10-13 min) it increases to 100% where it is kept for 3 min (13-16 min) and this was followed by an equilibration step from 18-21 min only with solvent A. A photo Diode-Array Detector (DAD) was used to detect analytes at 260nm, using a highly sensitive 60 mm flow cell. The separation between dead volume peak, cAMP monomer, linear AMP monomer, and oligomers was observed in HPLC

chromatogram. The dead volume peak components were characterized by mass spectrometry as open ring structures lacking phosphate moiety, i.e., adenosine and/or adenine. As shown in Figure 4.1, AMP (peak 1), ADP (peak 2) and ATP (peak 3) eluted at ~6.1 min, 6.6 min and 7.0 min, respectively. The peaks that eluted at later time points i.e., peaks 4 and 5 (after peak 3) contain more negative charge as compared to ATP (more than three phosphate groups), hence, indicating the potential formation of higher oligomers.
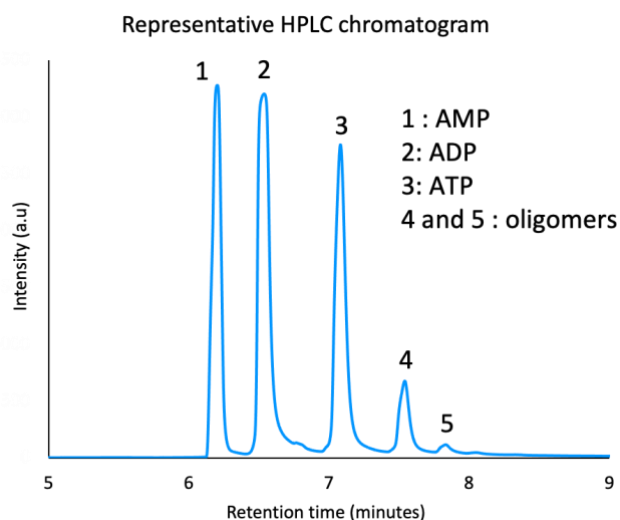

Figure 4.1: Representative High-Performance Liquid Chromatography (HPLC) trace showing the elution profile of various moieties (highlighted in the legend) present in the reaction mixture as separated on a DNAPac PA 200 column, using a gradient of sodium perchlorate (see methods for details). The y-axis depicts the absorbance at 260 nm and the x-axis shows a zoomed-in region of the retention time (in min), for better resolution between the different peaks/species eluting in the gradient.

4.2.2.4. Liquid Chromatography-Mass spectrometry (LC-MS) analysis

LC-MS analysis of the samples was performed on a Sciex X500R QTOF mass spectrometer (MS) fitted with an Exion-LC series Ultra Performance Liquid Chromatography (UPLC; Sciex), using an information-dependent acquisition (IDA) scanning method. The acquired data was analyzed using the Sciex OS software (Sciex; University of Florida). The crude reaction mixture was separated on Zorbax

C8 column (dimensions: 4.6×150 mm, 3.6 µm particle size) (Thermo Scientific), which was fitted with a guard column. A solvent system comprising of nanopure water containing 0.1% formic acid and acetonitrile containing 0.1% formic acid was used to separate the oligomers. The gradient used was: 0% solvent B for 3 minutes, an increase of solvent B from 0 to 20% was facilitated from 3 min to 7 min, and in next 9 min (i.e., 7-16 min) it increased to 30%, which is followed by a further increase to 100% from 16-18 min. The column was kept at 100% solvent B for the 3 min from 18-21 min, followed by equilibration with solvent A from 22-25 min. All the mass acquisition was performed using electron spray ionization (ESI) with the following parameters: turbo spray ion source, medium collision gas, curtain gas=30 L/min, ion spray voltage=5500 V (positive mode), at 500°C. TOF-MS acquisition was done at declustering potential of 80 V, with 20 V as spread and using 10 V collision energy. To perform TOF-MS–MS analysis, 50 V collision energy with a 20 V spread was used. As the mass acquisition was carried out in positive mode, the observed masses corresponded to the mass of the H+-adduct of the parent molecule. The presence of a specific species/molecule was confirmed by the presence of precursor mass within 5-ppm error range, as well as its fragmentation pattern. For relative quantification of different molecules, the yield fraction was calculated by dividing the area under its LC-MS peak by the sum total sum under all the peaks.

4.3. Results

4.3.1. Nonenzymatic oligomerization of 5′-adenosine triphosphates (ATP) under DH–RH conditions

We first investigated the enzyme-free oligomerization of ATP under both aqueous heating and dry heating control reaction conditions. In the aqueous heating reactions, 5 mM ATP solution (pH 7) was sealed and kept at 90°C. After varying time periods (0 hr, which is the initiation of the reaction), 1 hr, 4 hr, 8 hr and 12 hr), samples were withdrawn and analyzed using HPLC and LC-MS. Under dry heating conditions, 5 mM ATP solution (pH 7) was dried at 90°C under a gentle flow of nitrogen (N_2) (see

methods for details). After varying time periods, the dried sample was rehydrated and analyzed using HPLC and LC-MS.

As shown in Figure 4.2A, under aqueous heating conditions, only three peaks 1, 2 and 3 corresponding to AMP, ADP and ATP, respectively, were observed. With increasing incubation at 90°C, which is indicated by the various time points mentioned (Figure legend), a decrease in the intensity of peak 3 (ATP) and a corresponding increase in the intensities of peak 1 (AMP) and peak 2 (ADP) were observed depicting the hydrolysis of ATP to ADP and AMP under these conditions. In dry heating ATP reactions, with increasing incubation time, a decrease in ATP peak and an increase in the intensities of AMP and ADP peak was observed, which was similar to the aqueous heating reactions. Interestingly, new peaks which eluted at later retention times (peaks 4 and 5) were also observed within one hour indicating the potential presence of oligomers. These peaks were observed even after 8 hours of dry heating but disappeared after 12 hours. This could be because of the instability of resultant oligomers at 90°C (high temperature) for longer time periods.

The formation of oligomeric species was observed under dry heating conditions but not under aqueous heating conditions. This could be because of the competitive reaction of ATP hydrolysis (prevalent under aqueous heating conditions) vs its oligomerization. Under dry heating scenarios, the hydrolysis of ATP is reduced, due to the absence of water. Moreover, such conditions also enhance the proximity of molecules, thereby facilitating concentration-dependent reactions such as oligomerization reactions [14,27]. Although insightful, the presence of complete dry niches on the early Earth is improbable [5]. As alluded to earlier, DH-RH conditions are a recurrent phenomenon in terrestrial geothermal pools and are shown to facilitate the formation and condensation of a variety of molecules [28–30].

With this line of thought, we next investigated the enzyme-free oligomerization of ATP under early Earth simulated DH-RH conditions at pH 7 and 90°C. Upon analysis of various DH-RH cycles, the new peaks (peaks 3 and 4) corresponding to potential

90

oligomers were observed within 1 DH-RH cycle (Figure 4.3A). The intensity of potential oligomeric peaks increased after 1 DH-RH cycle and stayed up to the 3 DH-RH cycles. Encouraged by the presence of oligomeric species under DH-RH, we further wanted to optimize the conditions such as temperature and pH for these reactions. Towards this, we varied the pH and temperature of these reactions. For the pH standardization, three different pHs were investigated i.e., pH 2, pH 7 and pH 9, and samples from all three sets were analyzed after 5 DH-RH cycles (Figure 4.3B). At pH 7 and 9, oligomeric peaks (peaks 3 and 4) were observed after 5 DH-RH cycles. At pH 2, no visible oligomerization peaks (3 and 4) were observed, however, another peak eluted at around 1.2 minutes (breakdown; Figure 4.3B inset). Previous studies have confirmed that this peak is due to the acid-catalyzed depurination that occurs under these conditions [3,5]. Therefore, given the instability of glycosidic bond and phosphodiester bond under acidic and alkaline regimes, respectively, further oligomerization reactions were all performed at pH 7.
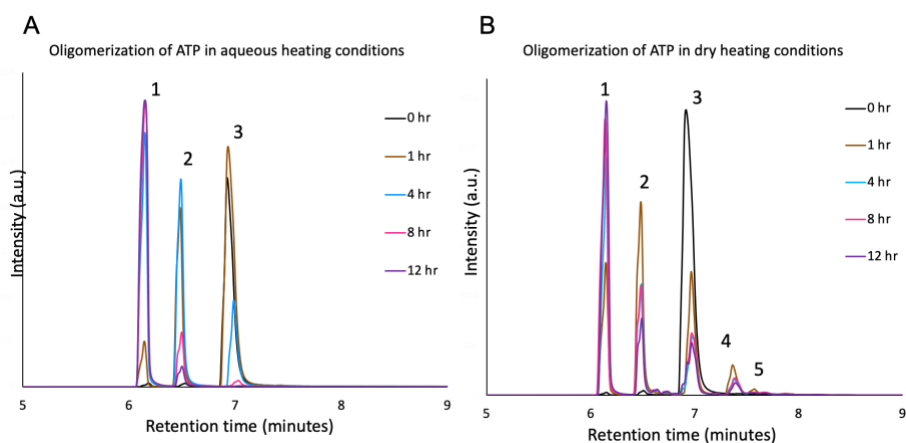

Figure 4.2: Nonenzymatic oligomerization of 5 mM ATP under aqueous heating reaction conditions (A) and dry heating reaction conditions (B) at 90°C and pH 7. HPLC chromatogram depicts different chemical moieties (1: AMP, 2: ADP, 3: ATP, 4 and 5: potential oligomers) after varying time periods, as highlighted in the legend. The y-axis and x-axis show the absorption at 260 nm and the retention time (in min), respectively. N = 3.

Next, to optimize the temperature, reactions were performed at two different temperatures i.e., 60°C and 90°C. The presence of oligomeric species was confirmed and quantified by LC-MS. Intact oligomers up to intact trimers and tetramers were observed at 60°C and 90°C, respectively (Figures 4.3C, 4.3D, 4.6A and Table 4.1).). Upon quantification, the yield fraction of the oligomers was observed to increase significantly upon increasing the number of DH-RH cycles in both the reaction sets. The fraction of oligomers was found to be significantly higher in the reactions performed at 90°C when compared to the reactions undertaken at 60°C. Henceforth, all the reactions were performed at 90°C and pH 7 and for 5 DH-RH cycles, after which samples were withdrawn and analyzed using LC-MS at the initiation of the reaction (cycle 0), after 1, 3 and 5 DH-RH cycles, respectively.
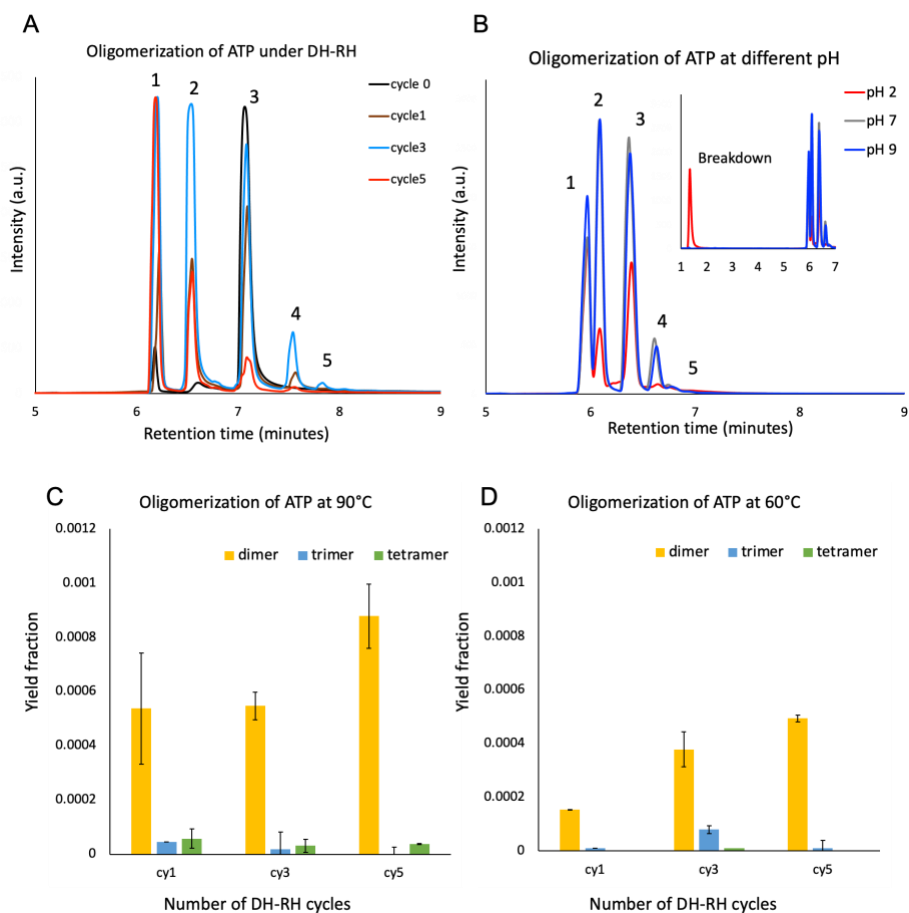
Figure 4.3: Nonenzymatic oligomerization of 5 mM ATP under DH-RH cycling conditions. A) HPLC chromatogram showing the presence of different chemical moieties (1: AMP, 2: ADP, 3: ATP, 4 and 5: potential oligomers) after various DH-RH cycles (0, 1, 3 and 5, as depicted in legend) at pH 7 and 90°C. B) HPLC chromatogram showing different chemical moieties after 5 DH-RH cycle at 90°C and different pH (2, 7 and 9 as shown in the legend). The y-axis and x-axis show the absorption at 260 nm and the retention time (min), respectively. N = 3. C) and D) show yield fraction of intact oligomers after varying DH-RH cycles (1, 3 and 5, respectively as indicated on x-axis) at 90°C and 60°C, respectively, and these were quantified using LC-MS.

Table 4.1: Masses observed during Liquid Chromatography-Mass spectrometry (LC-MS) characterization of nonenzymatic oligomerization of ATP, under DH-RH cycling conditions, along with calculated parts per million (ppm) error. Different oligomeric species for dimer, trimer and tetramer were observed as indicated below.

		z	m/z (calculated)	m/z (observed)	error (in ppm)
Monomer	ATP	1	508.0030	508.0031	0.26
	ADP	1	428.0367	428.0367	-0.05
	AMP	1	348.0704	348.0700	-0.86
Dimer	AMP-AMP	1	677.1229	677.1236	0.997
	AMP-ADP	1	757.0892	757.0891	-0.14
	AMP-ATP or ADP-ADP	1	837.0555	837.0546	-1.06
Trimer	AMP-AMP-ATP or AMP-ADP-ADP	1	1166.1080	1166.1061	-1.7
Tetramer	AMP-AMP-AMP-AMP	1	1335.2279	1335.2331	3.903
	AMP-AMP-AMP-ADP	1	1415.1942	1415.1902	-2.89
	AMP-AMP-AMP-ATP or AMP-AMP-ADP-ADP	1	1495.1606	1495.1632	1.779

4.3.2. Effect of different salts on the nonenzymatic oligomerization of ATP

Effect of cations

NTPs are inherently activated molecules, which recently have been shown to be plausible on prebiotic Earth[22]. As mentioned earlier, they are the biological substrates for extant polymerases used to replicate nucleic acid, hence, their prebiotic plausibility and enzyme-free oligomerization can potentially bridge the 'abiotic' to 'biotic' transition[23]. However, one of the major concerns of using ATP as a substrate for nonenzymatic reactions is its spontaneous hydrolysis to 5′-adenosine diphosphates (ADP) and 5′-adenosine monophosphates (AMP). Although AMP have been demonstrated to undergo enzyme-free oligomerization under acidic geothermal conditions[4], the systematic characterization of the resultant oligomers has shown the presence of abasic sites due to acid-catalyzed depurination[3,31]. This emphasized the requirement of intrinsically activated nucleotides such as ATP for enzyme-free oligomerization reactions.

Importantly, one of the salient characteristics of contemporary polymerases is the presence of Mg^{2+} ions in their active site[2,26]. It is noteworthy that divalent cations are shown to possess phosphoryl transferase and kinase activity[24,25,32]. Moreover, prebiotic soup in which processes that led to the emergence of life is considered to have been facilitated, is thought to be heterogeneous in nature containing several different molecules including various metal ions. Thus, the presence of different metal ions also has to be accounted for when simulating prebiotically pertinent reactions as this would provide realistic insights into how reactions like nonenzymatic oligomerization would have actually panned out. Therefore, we investigated the effect of different biologically relevant metal ions including Mg^{2+}, Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}, on the nonenzymatic oligomerization of ATP. As a first step, sulfate salts of all these metal ions were used in the reactions. Four different ratios of ATP to metal ions including 1:1, 1:2, 1:4 and 2:1, were investigated to optimize the metal-mediated oligomerization. All other reactions such as concentration of ATP,

temperature, pH and DH-RH cycles were kept the same as that of previously optimized conditions.

The presence of intact oligomers (different dimers viz. ATP-AMP or ADP-ADP, ADP-AMP and AMP-AMP) was analyzed and confirmed using LC-MS (Figure 4.5A). As shown in Figure 4.4, various types of intact dimers were observed in all the reactions. In the case of Mg^{2+} containing reactions, both 1:1 and 1:2 ratios showed the highest number of intact dimers (total number of green boxes), which were observed in both cycle 3 and cycle 5. In ATP: Mg^{2+}:: 1:4 and 2:1 reactions, less number of oligomers were observed. Also, in the case of ATP: Mg^{2+}:: 2:1 reaction, the oligomers formed by cycle 3 disappeared by cycle 5. In the cases of other metal ions i.e., Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}, relatively lesser number of oligomers were observed (lesser number of green boxes). Hence, Mg^{2+} was utilized for further oligomerization reactions. Importantly, since higher amount of divalent cations is also known to catalyze hydrolysis reactions, we chose to use ATP: Mg^{2+}::1:1 going forward.

Effect of anions

In order to investigate the effect of the anion component, the nonenzymatic oligomerization reactions were performed with two different salts of Mg^{2+} i.e., sulfate salt and chloride salt. Both the salts were used in a 1:1 ratio with respect to ATP (based on results from the aforementioned section). The reaction mixture was analyzed after various DH-RH cycles (cycle 0, cycle 1, cycle 3 and cycle 5) using HPLC and the presence of oligomeric species was confirmed using LC-MS analysis. As observed in Figure 4.5B, after 5 DH-RH cycles, oligomeric species were observed in the presence of both $MgSO_4$ and $MgCl_2$ salt. The extent of oligomerization was higher in $MgCl_2$-containing reactions when compared to $MgSO_4$-containing ones. Interestingly, the total oligomer yield (%) was higher for the $MgCl_2$-containing reactions when compared to the ATP only reactions (Figures 4.6B and 4.6C). However, oligomer only up to intact dimer(s) was observed in these $MgCl_2$-containing reactions when compared to intact tetramer in its absence.

Figure 4.4: Nonenzymatic oligomerization of ATP in the presence of different metal sulfate salts. The reactions were performed using 5 mM ATP at pH 7 and 90°C under DH-RH cycling conditions for various ATP:Metal ion (1:1, 1:2, 1:4 and 2:1). The presence of different types of intact dimer species including ATP-AMP, ADP-AMP and AMP-AMP, after different DH-RH cycles in the presence of various metal ions (Mg^{2+},

Legend:
- Present
- Absent
- 1 = AMP-ATP or ADP-ADP
- 2 = ADP-AMP
- 3 = AMP-AMP

Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}), has been indicated in the figure. Red color depicts absence and green color depicts presence of the corresponding entity.
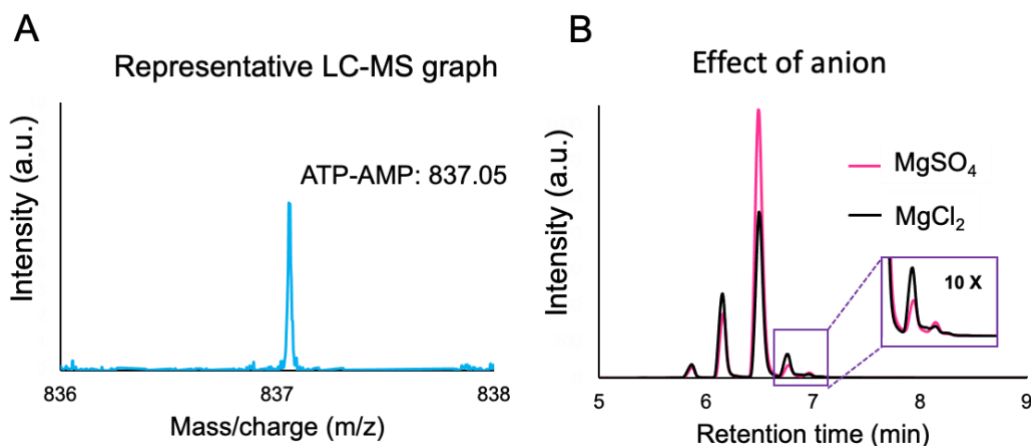
Figure 4.5: A) Representative LC-MS trace showing the presence of a ATP-AMP dimer with an m/z of 837.05. B) HPLC chromatogram showing different chemical moieties after 5 DH-RH cycles at 90°C and pH 7 in the presence of ATP:Mg^{2+}::1:1 for sulfate salt or chloride salt. The inset shows 10X zoomed version of the potential oligomers. The y-axis and x-axis show the absorption at 260 nm and the retention time (min), respectively. N = 3.
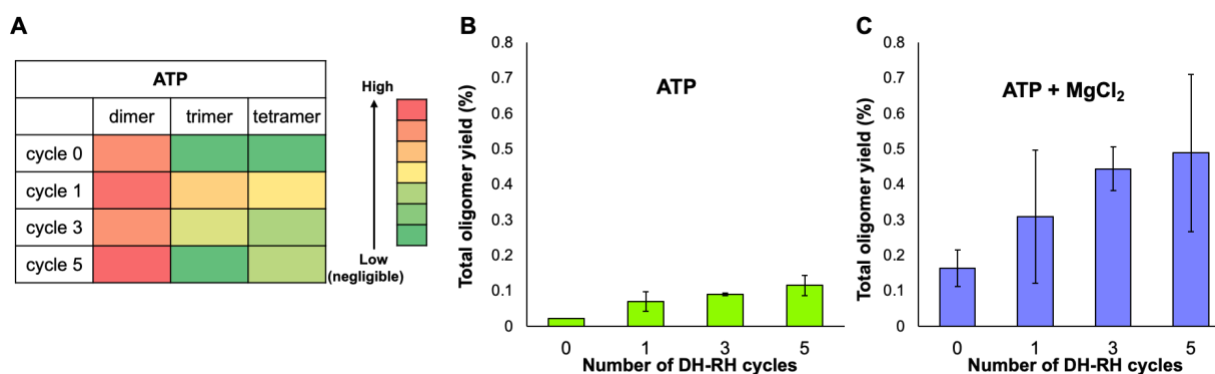
Figure 4.6: A) Heat map showing the distribution of oligomeric species (dimer, trimer and tetramer), after multiple DH-RH cycles using LC-MS. For quantification of the total yield, the sum of all the different oligomer species was used. B) and C) show the quantified total oligomeric yield (%) after various DH-RH cycles as indicated on the x-axis, for ATP and ATP:$MgCl_2$::1:1 reactions, respectively. N = 3, error bars = s.d.,

4.4. Discussion

Nonenzymatic oligomerization to form RNA is a crucial step in the emergence of a putative RNA World on the early Earth [1]. Towards this, most of the previous studies have demonstrated nonenzymatic oligomerization of imidazole-activated nucleotides (ImpN) [6–9]. Similar attempts to form oligomers with prebiotically relevant non-activated NMPs were observed to be facilitated by high temperature, acidic pH and DH-RH cycles [3,4,33]. However, the systematic characterization of the oligomers formed under these conditions were found to be abasic due to the cleavage of the glycosidic bond. Given this, we undertook studies on enzyme-free oligomerization of cyclic nucleotides (cNMPs) under multiple DH-RH cycles (Chapter 2). Herein, we have used ATP to investigate their oligomerization potential under early Earth conditions. As mentioned earlier, NTPs are the biological substrates for contemporary polymerases. Hence, enzyme-free oligomerization using NTPs can potentially bridge the 'abiotic' and 'biotic' RNA formation scenarios. Towards this, we explored several different parameters such as niche parameters (dry heating vs aqueous heating vs DH-RH cycles), ATP concentrations (5 mM and 10 mM), temperature (60°C and 90°C) and pH (2, 7 and 9), on the oligomerization of ATP.

As a first step, the nonenzymatic oligomerization of 5 mM ATP was investigated under dry heating and aqueous heating conditions. The reactions were performed at 90°C and pH 7 as ATP contains a pyrophosphate which can, in principle, act as a good leaving group similar to the imidazole in ImpNs. Therefore, the nonenzymatic oligomerization of ATP should not require extreme (acidic) pH conditions as that of NMPs. Under aqueous heating conditions, spontaneous degradation of ATP to ADP and AMP was observed. No oligomeric species were observed under aqueous heating conditions even after 12 hours of incubation at 90°C. In the case of dry heating, HPLC peaks corresponding to potential oligomers were observed within 1 hour of incubation and even after 12 hours. This could be because of increasing the effective concentration of ATP monomers under dry heating conditions, in turn increasing their proximity, resulting in uphill oligomerization reactions. Although insightful, the presence of completely dry niches on the early Earth is uncertain.

As the next logical step, nonenzymatic oligomerization reactions using ATP, were performed under prebiotically relevant DH-RH cycling conditions. As indicated earlier, the dry phase in a DH-RH cycle enhances the proximity of monomers, thus aiding oligomerization (similar to dry heating reactions), and the rehydrated phase facilitates the redistribution of molecules[15]. Subsequently, all the enzyme-free oligomerization of ATP was performed under DH-RH cycling conditions. Under these conditions, oligomers were observed within 3 cycles and persisted till 5 cycles. The oligomers were observed under pH 7 and pH 9, whereas at pH 2, a breakdown peak corresponding to adenine was observed. This could be because of the acid-catalyzed depurination of monomers under such conditions[3,34]. As alkaline pH is known to be detrimental to the phosphodiester bond, pH 7 was used for further reactions. Next, the DH-RH reactions were performed at this pH 7 and at two different temperatures i.e., 90°C and 60°C. At 60°C, up to 0.02% intact dimer was observed within cycle 1 (as confirmed by LC-MS) which enhanced to 0.05% after 5 DH-RH cycles. In the case of reactions at 90°C, up to 0.056% intact dimer was observed within cycle 1, which increased to 0.09% after 5 DH-RH cycles. Additionally, intact trimers and tetramers were also observed at 90°C after cycle 1, whereas at 60°C, intact trimer was observed only after cycle 3.

Next, we evaluated the effect of the presence of different metal salts as co-solutes on these reactions. Among all the cations investigated, the highest number of intact dimers were observed when Mg^{2+} (sulfate salt) was used (in ATP: Mg^{2+}::1:1 and 1:2 ratios), respectively. Upon varying the anion (chloride or sulfate), the extent of oligomerization was higher in $MgCl_2$-containing reactions when compared to the $MgSO_4$-containing ones. Interestingly, the oligomer yield (%) in $MgCl_2$-containing reactions was significantly higher as compared to just the ATP only reaction (Figure 4.6B and 4.6C). Nonetheless, the detailed chemical nature of the oligomers could not be ascertained as different types of oligomers, for example: ADP-ADP and AMP-ATP (in the case of dimers), can have the same m/z. However, as the reactions are occurring under repeated DH-RH cycles and in presence of Mg^{2+}, we hypothesize

that the Mg^{2+} can coordinate with pyrophosphate, making it a good leaving group. This would mean that, in principle, the formation of oligomers like AMP-ATP could be favoured more instead of ADP-ADP dimer under the aforesaid condition. These could not be fully characterised mainly due to yield related issues. Altogether, these results demonstrate nonenzymatic oligomerization of ATP under prebiotically relevant geothermal conditions. Additionally, the presence of oligomers was also demonstrated under a realistic scenario by using different metal salts as co-solutes, which would have been the case in a heterogeneous prebiotic soup. The presence of metal salts would have facilitated these oligomerization reactions under such scenarios, as was observed in the case of $MgCl_2$. This further emphasizes the importance of considering the effect of the co-solute molecules that would have been present in the surrounding environment. Thus, accounting for the heterogeneity inherent to the prebiotic soup could actually provide realistic insights about the formation of a putative RNA World on the early Earth.

4.5. References

1.   Gilbert, W. Origin of life: The RNA world. *Nature* **319**, (1986).

2.   Cramer, P. *et al.* Structure of eukaryotic RNA polymerases. *Annu Rev Biophys* **37**, 337–352 (2008).

3.   Mungi, C. V., Rajamani, S., Mungi, C. V. & Rajamani, S. Characterization of RNA-Like Oligomers from Lipid-Assisted Nonenzymatic Synthesis: Implications for Origin of Informational Molecules on Early Earth. *Life (Basel, Switzerland)* **5**, 65–84 (2015).

4.   Rajamani, S. *et al.* Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Orig. Life Evol. Biosph.* **38**, 57–74 (2008).

5.   Dagar, S., Sarkar, S. S. & Rajamani, S. Geochemical influences on nonenzymatic oligomerization of prebiotically relevant cyclic nucleotides. *RNA* **26**, 756–769 (2020).

6. Lohrmann, R. & Orgel, L. E. Prebiotic Activation Processes. *Nature* **244**, 418–419 (1973).

7. Blain, J. C. & Szostak, J. W. Progress Toward Synthetic Cells. *Annu. Rev. Biochem.* **83**, 615–640 (2014).

8. Puthenvedu, D., Majerfeld, I. & Yarus, M. Non-Watson-Crick RNA synthesis suited to origin functions. *RNA* **24**, 90–97 (2017).

9. Majerfeld, I., Puthenvedu, D. & Yarus, M. Cross-backbone templating; ribodinucleotides made on poly(C). *RNA* **22**, 397–407 (2016).

10. Yi, R., Hongo, Y. & Fahrenbach, A. C. Synthesis of imidazole-activated ribonucleotides using cyanogen chloride. *Chem. Commun.* **54**, 511–514 (2018).

11. Costanzo, G. *et al.* Generation of RNA Molecules by a Base-Catalysed Click-Like Reaction. *ChemBioChem* **13**, 999–1008 (2012).

12. Verlander, M. S. & Orgel, L. E. Analysis of High Molecular-Weight Material From Polymerization of Adenosine Cyclic 2',3'-Phosphate. *J. Mol. Evol.* **3**, 115–120 (1974).

13. Costanzo, G. *et al.* Non-Enzymatic Oligomerization of 3', 5' Cyclic AMP. *Plos One* **11**, e0165723 (2016).

14. Verlander, M. S., Lohrmann, R. & Orgel, L. E. Catalysts for the self-polymerization of adenosine cyclic 2',3'-phosphate. *J. Mol. Evol.* **2**, 303–316 (1973).

15. Damer, B. & Deamer, D. Coupled Phases and Combinatorial Selection in Fluctuating Hydrothermal Pools: A Scenario to Guide Experimental Approaches to the Origin of Cellular Life. *Life* **5**, 872–887 (2015).

16. Gulick, A. Phosphorus as a factor in the origin of life. *Am. Sci.* **43**, 479–489 (1955).

17. Schwartz, A. W. Phosphorus in prebiotic chemistry. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 1743–1749 (2006).

18. Pasek, M. A., Dworkin, J. P. & Lauretta, D. S. A radical pathway for organic phosphorylation during schreibersite corrosion with implications for the origin of

life. *Geochim. Cosmochim. Acta* **71**, 1721–1736 (2007).

19.   Pasek, M. A. Schreibersite on the early earth: Scenarios for prebiotic phosphorylation. *Geosci. Front.* **8**, 329–335 (2016).

20.   Pasek, M. a & Lauretta, D. S. Aqueous corrosion of phosphide minerals from iron meteorites: a highly reactive source of prebiotic phosphorus on the surface of the early Earth. *Astrobiology* **5**, 515–35 (2005).

21.   Toner, J. D. & Catling, D. C. A carbonate-rich lake solution to the phosphate problem of the origin of life. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 883–888 (2020).

22.   Kim, H. J. & Benner, S. A. Abiotic Synthesis of Nucleoside 5′-Triphosphates with Nickel Borate and Cyclic Trimetaphosphate (CTMP). *Astrobiology* **21**, 298–306 (2021).

23.   Lin, H., Jiménez, E. I., Arriola, J. T., Müller, U. F. & Krishnamurthy, R. Concurrent Prebiotic Formation of Nucleoside-Amidophosphates and Nucleoside-Triphosphates Potentiates Transition from Abiotic to Biotic Polymerization. *Angew. Chemie - Int. Ed.* **30332**, (2021).

24.   Sträter, N. & Lipscomb, W. N. Two-Metal Ion Catalysis in Enzymatic Acyl-and Phosphoryl-Transfer Reactions. *Angew. Chemie (International Ed. English)* **35**, 2024–2055 (1996).

25.   Sigel, H. & Tribolet, R. Synergism between different metal ions in the dephosphorylation of adenosine 5'-triphosphate (ATP) in mixed metal ion/ATP systems, and influence of a decreasing solvent polarity (dioxane-water mixtures) on the dephosphorylation rate. Effects of Mg2+, Na. *J. Inorg. Biochem.* **40**, 163–179 (1990).

26.   Doublie, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2Å resolution. *Nature* **39**, 251–258 (1998).

27.   Tapiero, C. M. & Nagyvary, J. Prebiotic Formation of Cytidine N ucleotides. *Nature* **231**, 42–43 (1971).

28. Becker, S. *et al.* Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat. Commun.* **9**, 1–9 (2018).

29. Becker, S. *et al.* Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science (80-.).* **82**, 76–82 (2019).

30. Forsythe, J. G. *et al.* Ester-Mediated Amide Bond Formation Driven by Wet-Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. *Angew. Chemie - Int. Ed.* **54**, 9871–9875 (2015).

31. Mungi, C. V, Bapat, N. V, Hongo, Y. & Rajamani, S. Formation of Abasic Oligomers in Nonenzymatic Polymerization of Canonical Nucleotides. *Life* 1–11 (2019).

32. Gopinath, P., Ramalingam, V. & Breslow, R. Magnesium pyrophosphates in enzyme mimics of nucleotide synthases and kinases and in their prebiotic chemistry. *Proc. Natl. Acad. Sci.* **112**, 12011–12014 (2015).

33. Olasagasti, F. & Rajamani, S. Lipid-assisted polymerization of nucleotides. *Life* **9**, 1–10 (2019).

34. An, R. *et al.* Non-Enzymatic Depurination of Nucleic Acids: Factors and Mechanisms. *PLoS One* **9**, e115950 (2014).

Chapter 5

Porphyrin in Prebiotic Catalysis: Ascertaining a Route for the Emergence of Early Metalloporphyrins

(Adapted from, Dagar et.al. 2022; ChemBioChem)

5.1. Introduction

Electron transfer reactions are ubiquitously utilized in contemporary biology to maintain chemical disequilibrium and to produce energy. Such redox reactions are predominantly catalyzed by metalloenzymes, which constitute about one-third of the contemporary enzyme repertoire[1,2]. The metal ions present in these enzymes are either part of the catalytic core, or the structural scaffold[3–6]. Previous reports have demonstrated metal and mineral-mediated catalysis in prebiotic chemistry[7–13], especially in the context of Iron-sulfur World hypothesis[5,14,15]. Iron-sulfur clusters and metal-coordinated porphyrins are potentially few of the earliest catalysts to have emerged on the prebiotic Earth owing to their presence in the catalytic core of some of the most ancient enzymes[1,2,16,17]. In this regard, few earlier studies have shown the interaction of such iron-sulfur clusters with peptide chains to form a protoenzyme-like entity[18,19]. In 2018, Bonfio et. al. demonstrated the ability of iron-coordinated peptide scaffolds to transfer electrons and generate a pH gradient across membranes[15]. These studies have also tried to address the putative evolution of this class of protoenzymes from free metal ions, under early Earth conditions. Nonetheless, evolutionary studies in the context of metal-organic scaffolds such as metalloporphyrins are still lacking.

Metalloporphyrins are utilized by a wide variety of extant enzymes including magnesium porphyrin in chlorophyll pigments, heme (ferric porphyrin) in hemoglobin, peroxidase and myoglobin, and copper porphyrin in cytochrome c oxidase, to name a few [20,21]. A highly desired property of transition metal ions is their ability to attain multiple oxidation states, which is crucial to catalyze different redox reactions. This "tunability" is employed in biology by coordinating metal ions with the porphyrin rings. Given this, porphyrin has been studied previously as a modular catalyst in biomimetic systems[20,22,23]. A pioneering study by Melvin Calvin noted a drastic increase in the catalytic efficiency of ferric ions when it was coordinated to a porphyrin scaffold[24,25]. This indicated a central role for metal-bound organic scaffolds such as porphyrins in modulating the catalytic

efficiency of metal ions. Recent studies have reported the prebiotically plausible abiotic synthesis of porphyrins, while they have also been detected in the interstellar dust[22,26–32]. Considering their prebiotic plausibility, presence in contemporary biology and the functional diversity of porphyrins, it is reasonable to hypothesize that porphyrins might have played a pivotal role in the emergence and evolution of metalloenzymes on the early Earth.

In this backdrop, we set out to investigate the role of porphyrin scaffolds on the oxidizing ability of metal ions potentially occurring in prebiotic Earth. Towards this, we first evaluated the oxidizing ability of different biologically relevant metal ions namely, magnesium, manganese, iron, cobalt, nickel, copper and zinc. Thereafter, the modulation of the oxidizing ability of the metal ions was studied when porphyrin was added as a co-solute in the reaction. We also investigated the ability of porphyrin to coordinate with these different metal ions under pertinent prebiotic settings. In this regard, the formation kinetics of different metalloporphyrins were studied using steady-state fluorescence and UV spectroscopy. Subsequently, the effect of porphyrin coordination on the oxidizing ability of metal ions was also investigated, to understand the oxidizing ability of preformed metal-coordinated porphyrin complexes.

We observed varying oxidizing ability for different metal ions, with Fe^{3+} showing the maximum oxidation among all the metal ions investigated. Interestingly, in co-solute reactions, significant changes were observed in the oxidizing ability of some metals, when compared to their respective free metal ion-based reactions, indicating the influence of porphyrins on the catalytic ability of these particular metal ions. We also report the formation of metal-coordinated porphyrin complexes under laboratory simulated prebiotic settings. The rate of coordination with the porphyrin was found to greatly vary for different metal ions. Also, the oxidizing ability of the preformed metal-coordinated porphyrins showed that the coordination process does not always result in an increase in the oxidizing ability of these metal ions. This is in contrast to a previously reported study wherein the catalytic efficiency was hypothesized to increase upon complexation[24,25]. Specifically, we observed that upon coordination, certain metal ions

showed a steep increase in catalysis (in the case of Co^{2+}), whereas in other cases, the coordination process dampens the catalytic efficiency of the otherwise catalytic metal ion (in the case of Fe^{3+}). Pertinently, we also report an alternate route by which porphyrins positively influenced the catalytic efficiency of Fe^{3+} ions that do not involve coordination. This was found to be achieved by forming non-coordinated assemblies between the metal ion and the porphyrin scaffold. Importantly, metal-coordinated porphyrin complexes could also catalyze the oxidation of nicotinamide adenine dinucleotide hydride (NADH), another biologically relevant molecule, essentially emphasizing their substrate versatility. Overall, our study highlights the important role that the porphyrin scaffold would have played in modulating the oxidizing ability of different metal ions. With this study, our attempt is to also set the stage for further exploration of porphyrin-based complexes and, in general, metal-organic scaffolds, for their potential to have played a central role in prebiotic catalysis.

5.2. Materials and Methods

5.2.1. Materials

All reagents were from Sigma Aldrich and used without further purification. The C18 chromatography column was purchased from Agilent Technologies (Santa Clara, CA, USA). All the oxidation reactions were performed in the anaerobic glove bag from Coy Labs under 95% N_2 and 5% H_2 (usually used to maintain anoxygenic conditions with oxygen levels < 200 ppm) environment. Nanopure water was filtered through a 0.22-micron filter, degassed and purged with N_2 to remove oxygen. The samples were transferred to glass vials covered with caps with air tight septum for HPLC-based analysis.

5.2.2. Methods
5.2.2.1. Experimental design
Hydroquinone (HQ) was used as the primary substrate for the catalytic reaction. This was used as a proxy for ubiquinol (UbQOH), which gets oxidized to ubiquinone [21]

(UbQone, Figure 5.1A) by cytochrome in extant biology. NADH oxidation was also evaluated to investigate the promiscuity of the metal-porphyrin complexes.

HQ gets oxidized to benzoquinone (BQ) as is shown in Figure 5.1B. This transition from HQ to BQ involves oxidation of the hydroxyl groups to keto groups, with an overall removal of 2 electrons. As a first step, the oxidizing ability of different metal ions in this process was evaluated. A typical reaction mixture contained 0.3 mM HQ with 0.3 mM metal ions (1:1 molar ratio, unless specified otherwise). All the oxidation reactions were performed under anoxygenic conditions in an anaerobic glove bag to mimic early Earth conditions [33]. The mixtures were incubated at 40°C and rotated at 350 rpm for four hours, without purging with inert gas or stirring (see Methods section for details). All the reactions were performed in unbuffered conditions using nanopure water (of resistivity 18 MΩ). The initial pH was near neutral for most of the reactions (except Fe^{3+} and Cu^{2+} containing reactions). The pH of the reaction mixture was also measured at the end of the reaction (i.e., after four hours), to check for any changes. The extent of oxidation at the initiation of the reaction, and after four hours, was analyzed using High Performance Liquid Chromatography (HPLC) (see Methods). HQ oxidation was quantified by using a standard graph obtained for various concentrations of HQ and BQ, respectively. The ratio of HQ to metal ions and the duration of the reaction were chosen based on the results we got for Cu^{2+}-mediated oxidation reactions that were performed in varying ratios (Figure 5.2).

Following this, the metal ion-mediated oxidation was performed in the presence of tetraphenyl porphyrin tetra sulfonic acid (TPPS) as a co-solute. In these co-solute reactions, in addition to 0.3 mM HQ and 0.3 mM metal ion, 0.03 mM TPPS (ten-fold lesser in concentration than the HQ and metal ions) was also added.

Subsequently, the tendency of the different metal ions to form metal-TPPS complexes (M-TPPS) under prebiotically relevant conditions was investigated by incubating 0.3 mM of the respective metal ions with 0.03 mM TPPS at high temperature (70°C). The formation of M-TPPS was monitored using UV spectroscopy, steady-state fluorescence spectroscopy and High-resolution Mass Spectrometry (HRMS).

As the next step, the oxidizing capability of the preformed M-TPPS complexes (metalloporphyrins) was evaluated. Towards this, M-TPPS were prepared by incubating 0.03 mM TPPS with 0.03 mM respective metal ion (1:1) at 70°C and followed till the completion of coordination. These complexes, which were devoid of any free metal ions, were then used to investigate the catalytic ability of preformed M-TPPS. In the case of Fe^{3+}-TPPS, a commercially acquired reagent was used as Fe^{3+}-TPPS did not form readily under our reaction conditions. For M-TPPS mediated HQ oxidation, 0.3 mM HQ was incubated with 0.03 mM M-TPPS (ten-fold lesser than HQ), at 40°C. In the case of M-TPPS mediated NADH oxidation reactions, 0.1 mM NADH was incubated at 25°C with 0.01 mM M-TPPS for varying time periods, and the reaction progress was analyzed with UV spectroscopy (see Methods).

5.2.2.2. High-performance liquid chromatography (HPLC)

HPLC analysis was performed using a Shimadzu HPLC system (CBM-20 A, quaternary pump LC-20AD; Shimadzu Corporations, Kyoto, Japan). The analytes were chromatographically resolved in the reverse-phase (Zorbax Eclipse Plus C18 column; 4.6 X 150 mm, 3.5 µm particle size, Agilent). For separation of Hydroquinone (HQ)/ Benzoquinone (BQ), a gradient of methanol was used as the eluting solvent. The HPLC run started with 100% nanopure water and 0% methanol and at a flow rate of 1 ml/ min. The methanol was increased to 15% over 15 min and was kept at this percentage until 18 min. Following this, it was reduced to 0% by 19 min and the column was equilibrated at this concentration for 2 more minutes. The analytes were detected using a photo Diode-Array Detector (DAD). A standard solution of 0.2 mM HQ was eluted at 5.5 minutes and 0.2 mM BQ eluted at 10.03 and 10.6 minutes, respectively. Typically, 30 µl of sample was withdrawn after different time periods, of which 25 µl was loaded onto the HPLC. Standard curves obtained for HQ (at 288 nm) and BQ (at 244 nm) were used to quantify the extent of oxidation by calculating the percentage of BQ produced in the reactions. The BQ was quantified by integrating the corresponding HPLC peaks. For TPPS complex containing reactions, in addition to the 21 minutes method, 100% methanol was run from 22 min to 29 min followed by equilibration with nanopure water from 30 min to 33 min.

5.2.2.3. Oxidation reactions with metal salts

An aqueous solution of each metal sulfate salt (20 mM for $Fe_2(SO_4)_3$ and 40 mM for other salts) was prepared. Unless otherwise reported, a typical reaction contained 0.3 mM hydroquinone (HQ) and 0.3 mM metal ion. For Fe^{3+}, one molecule of $Fe_2(SO_4)_3$ contains two atoms of Fe per formula unit. Therefore, in these reactions, 0.3 mM and 0.15 mM of sulfate salt concentration was used for to keep the molar ratio of HQ and metal ions constant as 1:2 and 1:1, respectively. The reaction mixture was then incubated at 40°C for four hours under anoxygenic conditions ($O_2 < 100$ parts per million (ppm)). 30 µl of sample was withdrawn immediately after the addition of metal salt in the reaction mixture (0 hr.) and then again after four hours of incubation (4 hrs.), out of which 25 µl was loaded onto the column for HPLC-based analysis.

5.2.2.4. Oxidation reactions with metal salts and TPPS as co-solute

A typical reaction contained 0.3 mM HQ, 0.3 mM metal ion and 0.03 mM TPPS (1:1:0.1 molecular ratio). Control reactions with only HQ or containing only TPPS along with HQ were also performed. Other conditions such as incubation at 40°C for four hours were kept constant, following which the samples were analyzed using HPLC by loading 25µl of the respective sample at a given time.

5.2.2.5. UV-vis spectroscopy

A UV-1800 Shimadzu spectrophotometer was used to measure UV-vis absorption spectra (scan range from 350 nm to 700 nm, interval = 1 nm) for the samples after various times. UV-vis spectrum was recorded at room temperature.

5.2.2.6. Steady-state fluorescence spectroscopy

A Fluoromax-4 spectrofluorometer (Horiba Scientific, Kyoto, Japan) was used for fluorescence measurements i.e., fluorescence quenching in the kinetics experiments and scattering in aggregation studies. For fluorescence-based assays, light with a wavelength of 414 nm was used for excitation. The excitation and emission slits were fixed at 2 nm and 1 nm, respectively. The fluorescence intensity for the emission signals

of a given sample was then scanned from 500 to 800 nm with 1 nm interval at 70°C, after different time periods of incubation.

For fast-kinetics experiments (in the case of Zn^{2+}, Cu^{2+} and Co^{2+}), the fluorescence intensity was measured at 693 nm for 0.03 mM TPPS solution at 70°C, till 300 seconds, after which 0.3 mM of corresponding metal ion was added and mixed. Following this, the fluorescence was monitored till 1500 seconds.

5.2.2.7. Rayleigh scattering to study aggregation

The light with a wavelength 414 nm was used for excitation and emission was also measured at 414 nm. The excitation and emissions slits were fixed at 2 nm. The fluorescence intensity was measured for 0.03 mM TPPS solution till 300 seconds, after which 0.3 mM of corresponding metal ions was added and mixed at room temperature. Following this, the fluorescence was then monitored till 1500 seconds.

5.2.2.8. Field-emission scanning electron microscopy (FE-SEM)

FE-SEM images were recorded using Zeiss Ultra Plus scanning electron microscope. Approximately 2.5 μl of the sample composed of 5 μM TPPS and 25 μM of $Fe_2(SO4)_3$ (1:10 of TPPS: Fe^{3+} ion) was prepared by drop casting on silicon wafers.

5.2.2.9. Differential Interference Contrast (DIC) Microscopy

Microscopic analysis of TPPS**Fe^{3+} non-coordinated aggregates was done using a DIC microscope AxioImager Z1 (Carl Zeiss, Germany), (NA = 0.75) under 40X objective. Approximately 10 μl mixture of 0.03 mM TPPS and 0.15 mM $Fe_2(SO4)_3$ (1:10 of TPPS: Fe^{3+} ion) was spread on a glass slide and covered by an 18X18 glass coverslip. Followed by which the sides of the cover slip were sealed with liquid paraffin and was observed under microscope.

5.2.2.10. Competition experiments for the formation of M-TPPS

0.03 mM TPPS was incubated with different mixtures of metal ions at 70°C and rotated at 350 rpm. The concentration of each metal ion was 0.03 mM. Four different combinations of different metal ions were used. Set 1 contained Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+},

Cu^{2+} and Zn^{2+}. Set 2 contained Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+} and Zn^{2+}. Set 3 contained Mn^{2+}, Fe^{3+}, Co^{2+} and Ni^{2+}. Set 4 contained Mn^{2+}, Fe^{3+} and Ni^{2+}. UV-vis absorption spectra of each of these combinations were recorded after different time periods to monitor the formation of different M-TPPS complexes.

5.2.2.11. Oxidation reactions with metal-coordinated TPPS (M-TPPS) complexes

The reaction mixture contained 0.3 mM HQ and 0.03 mM (1:0.1 atomic ratio) of corresponding M-TPPS complexes. All other reaction conditions such as incubation at 40°C for four hours, were kept the same as mentioned above. The reactions were then analyzed using HPLC. Turn over number (TON) were calculated by the following formula:

$$TON = \frac{\text{moles of BQ produced}}{\text{moles of M-TPPS used}} = \frac{\text{moles of HQ oxidized}}{\text{moles of M-TPPS used}} \qquad \text{Eq. 1}$$

$$\text{Moles of M-TPPS} = (\text{concentration of M-TPPS} * \text{Volume}) \qquad \text{Eq. 2}$$

$$\text{concentration of M-TPPS} = \frac{\text{concentration of HQ}}{10} \qquad \text{Eq. 3}$$

$$\text{Moles of M-TPPS} = \frac{(\text{concentration of HQ at t=0}) * \text{volume}}{10} \qquad \text{Eq. 4}$$

$$\text{Moles of HQ oxidized} = (\text{yield (\%) of HQ at t} = 0) * \text{volume} \qquad \text{Eq. 5}$$

Substituting Equations 4 & 5 in 1

$$TON = \big(\text{yield (\%)}\big)/10 \qquad \text{Eq.6}$$

5.2.2.12. M-TPPS mediated oxidation reactions of NADH

The reaction mixture contained 0.1 mM NADH and 0.01 mM (10: 1 molar ratio) of M-TPPS complexes. The mixture was then incubated at 25°C for four hours under anoxygenic condition ($O_2 < 200$ ppm). The reaction was analyzed using UV spectrophotometer by monitoring the disappearance of the signal at 340 nm, which corresponds to NADH absorption. Control reactions were performed i.e., spontaneous oxidation containing only NADH and reactions containing only Co^{2+} or only TPPS.

112

5.3. Results:

5.3.1. The oxidizing capability of metal ions

Firstly, we compared the oxidizing ability of different metal ions including Mg^{2+}, Mn^{2+}, Fe^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+} ions. These metal ions were selected based on their presence in the active core of extant metalloporphyrins [34–36]. On the prebiotic Earth, Fe^{2+} would have readily photo-oxidized to Fe^{3+} in the presence of UV and under aqueous conditions, irrespective of the presence of oxygen [37]. Therefore, owing to the susceptibility of Fe^{2+} towards oxidation, and the prebiotic relevance of Fe^{3+}, we evaluated the oxidizing ability of iron in its +2 as well as +3 states.
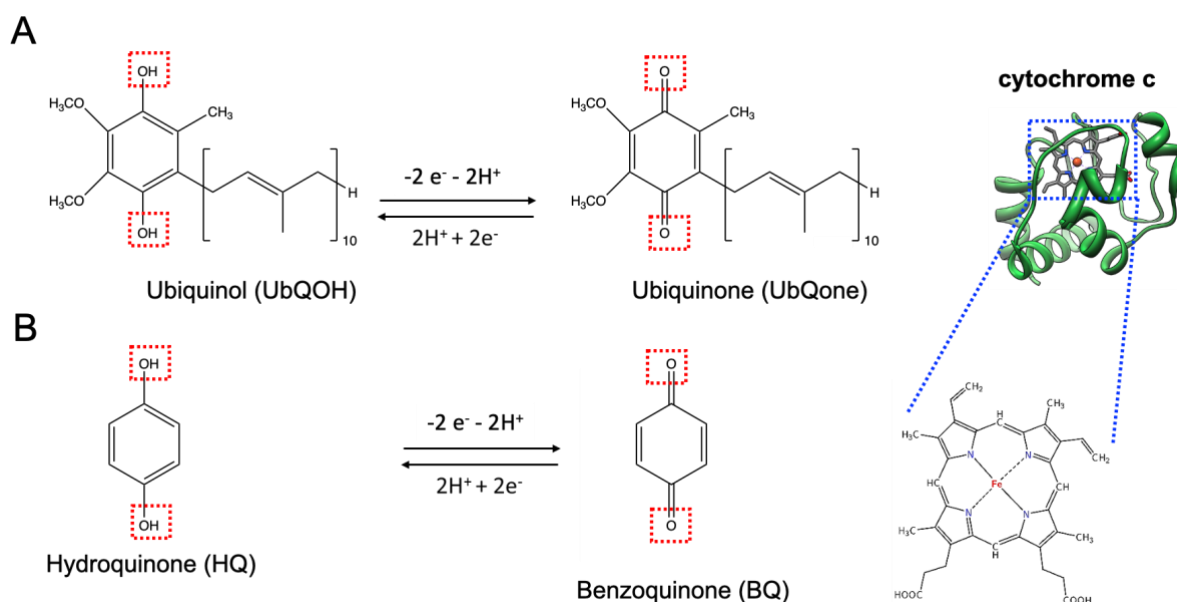

Figure 5.1: Schematic showing the oxidation of a quinol to its corresponding quinone. A) Oxidation of hydroxyl groups of ubiquinol (UbQOH) to keto groups to form ubiquinone (UbQnone), which involves the loss of 2 electrons and is catalyzed by the iron-porphyrin containing protein cytochrome c (shown in the upper right section; inset shows the iron-porphyrin complex). B) Oxidation of hydroquinone (HQ) to benzoquinone (BQ), a reaction used as a proxy to mimic oxidation of UbQOH to UbQnone. As shown, it also involves the oxidation of 2 hydroxyl groups by loss of 2 electrons, to result in keto groups in BQ.
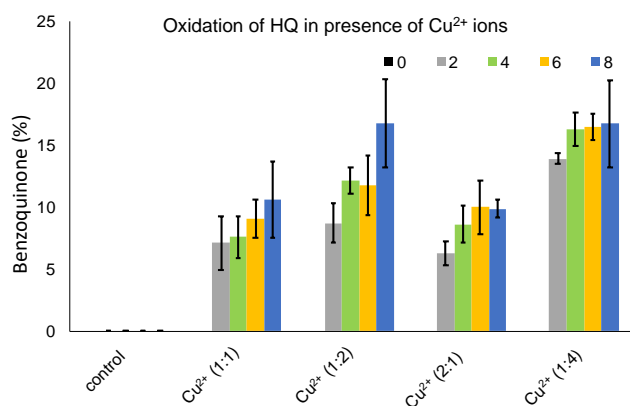
Figure 5.2: The graph shows the percentage of BQ produced in CuSO$_4$ containing reactions. The variations evaluated include no metal control, and varying ratios of HQ: Cu^{2+} i.e., 1:1 (Cu^{2+} (1:1)), 1:2 (Cu^{2+} (1:2)), 2:1 (Cu^{2+} (2:1)) and 1:4 (Cu^{2+} (1:4)), as depicted on X-axis. Y-axis shows the BQ (%) produced. Different colors indicate the different time points after which samples were retrieved i.e., immediately after salt addition (0), after 2hrs (2), 4hrs (4), 6hrs (6) and 8hrs (8), respectively. Error bars depict standard deviation; N=3.
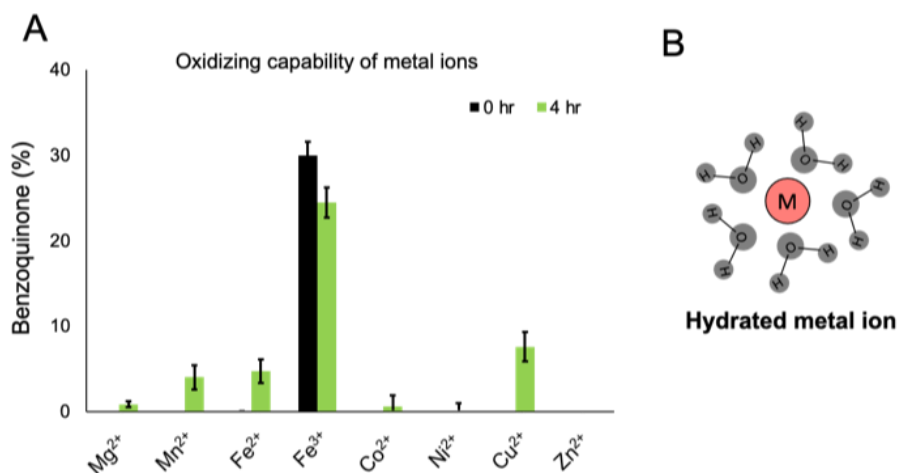

Figure 5.3: Oxidation of HQ to BQ in the presence of different metal ions. A) The bar graph shows the percentage of benzoquinone (BQ) produced from the oxidation of hydroquinone (HQ), with different metal ions used in 1:1 ratio of HQ to metal ions. Y-axis shows the percentage of benzoquinone (BQ) produced and X-axis shows the different metal ions used. Black and green colors indicate % BQ produced at the initiation (0 hr) and at the end (4 hr) of the reaction. Error bars depict standard deviation; N=3 (chemical replicates). B) Illustration depicting hydrated metal ions used in this reaction.

Table 5.1: Reduction potentials of the investigated metal ions and their respective oxidation state change (adapted from a:[38], b:[39–41]).

Metals	Standard reduction potential (Volts)	Changing oxidation state
Benzoquinone	0.643^a	$2\ e^-/2H^+$
	0.099^a	$1\ e^-$
Mg	-2.356^b	+2 to 0
Fe	$+0.771^b$	+3 to +2
	-0.44^b	+2 to 0
Co	-0.277^b	+2 to 0
Ni	-0.257^b	+2 to 0
Cu	0.3419^b	+2 to 0
	0.159^b	+2 to +1
	0.520^b	+1 to 0
Zn	-0.762^b	+2 to 0
Mn	-1.170^b	+2 to 0

No significant oxidation of HQ was observed in the control reaction, i.e., in the absence of any metal ions (Figure 5.3). Also, no oxidation was observed in the presence of Mg^{2+} and Zn^{2+} ions as these are redox inactive. In the presence of other metal ions, except for Fe^{3+}, no oxidation was observed at the initiation of the reaction. For Co^{2+} and Ni^{2+} ions, no significant oxidation of HQ was observed even after four hours (Figure 5.3). In the presence of Mn^{2+}, Fe^{2+} and Cu^{2+} ions, 4%, 4.7% and 7% of HQ oxidation, respectively, was observed after four hours. Interestingly, in Fe^{3+} containing reactions, up to 30% BQ was produced immediately upon the addition of the Fe^{3+} ions to the solution. The extent of oxidation by Fe^{3+} did not change significantly after four hours. Overall, the extent of oxidation for the different metal ions was in the following order: $Fe^{3+}> Cu^{2+}> Fe^{2+}> Mn^{2+}$. Interestingly, this trend of HQ oxidation remained unchanged even upon doubling the metal ion concentration (HQ: metal ion = 1:2)

The standard reduction potential associated with the oxidation of HQ to BQ for 2 e⁻/ 2 H⁺ and 1 e⁻/ 1 H⁺ is 0.643 V and 0.099 V, respectively [38]. In order to oxidize HQ, the reduction potential of the oxidizing agent (metal ions) needs to be higher (more positive) than the reduction potential of BQ. The standard reduction potentials for the transition of Fe^{3+} to Fe^{2+} and Cu^{2+} to Cu^{1+} are 0.771 V and 0.159 V, respectively [39–41] (Table 5.1), potentially justifying the ability of Fe^{3+} and Cu^{2+} to oxidize HQ presumably by 1 e⁻/ 1 H⁺ transfer process. Further, the slight oxidation observed in Fe^{2+} (-0.44 V) and Mn^{2+} (-1.170 V) could be attributed to their oxidation ability under aqueous conditions. Altogether, the ability of even free metal ions to oxidize HQ, essentially emphasized the influence of these ions on these prebiotically pertinent reactions, prior to the emergence of more efficient and complex (proto) metalloenzymes [7,24,25,42–45].

5.3.2. Reactions containing different metal ions along with TPPS as a co-solute

Taking into account the presumably heterogeneous nature of prebiotic soup, it is reasonable to assume that different metal ions would have co-existed with different co-solutes, including scaffolds similar to porphyrins [46]. And, the interactions between metal ions and such scaffolds might have impinged on their catalytic ability. In order to study this, we evaluated the effect of TPPS (tetraphenyl porphyrin tetra sulfonic acid) scaffold as a co-solute, by including them in the oxidation reaction mixtures containing the different metal ions, as depicted in Figure 5.4B (see Methods).

As seen in Figure 5.4A, negligible oxidation of HQ was observed in the control reactions containing just HQ and HQ with TPPS when there were no metal ions present. In the 'co-solute reactions', the oxidizing ability of Mg^{2+}, Ni^{2+}, Fe^{2+} and Zn^{2+} ions were found to be unaltered (Figure 5.4A). Interestingly, we observed significantly enhanced activity for Co^{2+} and Fe^{3+} ions in the presence of TPPS as a co-solute, as shown in Figure 5.4A. In the case of Co^{2+} and TPPS based co-solute reaction, up to ~12.3% of BQ was produced. Whereas, in the case of only free Co^{2+} ions, and only

TPPS containing reaction, negligible HQ oxidation was observed (Figure 5.3 and Figure 5.4 panel A, respectively). In the Fe^{3+} and TPPS based co-solute reaction, the HQ oxidation was enhanced to ~51.7% just at the initiation of the reaction. Whereas, in the case of only Fe^{3+} ions, up to 30% HQ oxidation was observed. Interestingly, in the case of Mn^{2+} and Cu^{2+} containing TPPS co-solute reactions, the BQ production was found to be diminished when compared to only Mn^{2+} ions (~4%) and Cu^{2+} ions (~7.6%) containing reactions (Figure 5.3, Figure 5.4 panel A). Importantly, a color change (from green to red) was also observed in co-solute reactions containing Fe^{3+}, Co^{2+}, Cu^{2+} and Zn^{2+} ions, over the course of the reaction.

Next, in order to investigate the underlying molecular change, UV absorbance spectroscopy was used to understand the change in the color and oxidation activity for Fe^{3+}, Co^{2+} and Cu^{2+} ions in the presence of TPPS as a co-solute. Free (uncomplexed) TPPS has an absorption maximum ~414 nm due to the transition of electrons from the ground state (S_0) to second excited state (S_2), along with four weak Q bands in the 500-700 nm region (due to S_0 to S_1 transitions) [47] (Figure 5.4 panel C; left graph). A change in the λ_{max} of TPPS was observed for the Co^{2+} (to 427 nm), Cu^{2+} (to 412 nm), Zn^{2+} (to 421 nm) and Fe^{3+} (to 432 and 493 nm) containing co-solute reactions. Additionally, a decrease in the number of Q-bands
in the 500-700 nm region was also observed, which is also reflected in the color change for all the aforementioned reactions (Figure 5.4 panel C; right graph shows this change for Cu^{2+}). Similar changes in the UV spectra of TPPS were also observed in the control reactions where it was incubated with Co^{2+}, Cu^{2+}, Zn^{2+} and Fe^{3+} ions, in the absence of HQ. All these observations indicated that the change in the activity of specific metal ions observed was potentially due to the coordination between metal ions and TPPS. This also implied that for specific metal ions, the coordination with TPPS occurs readily at 40°C (as reflected in the change in the UV spectrum) [48,49], which resulted in a change in their catalytic activity.
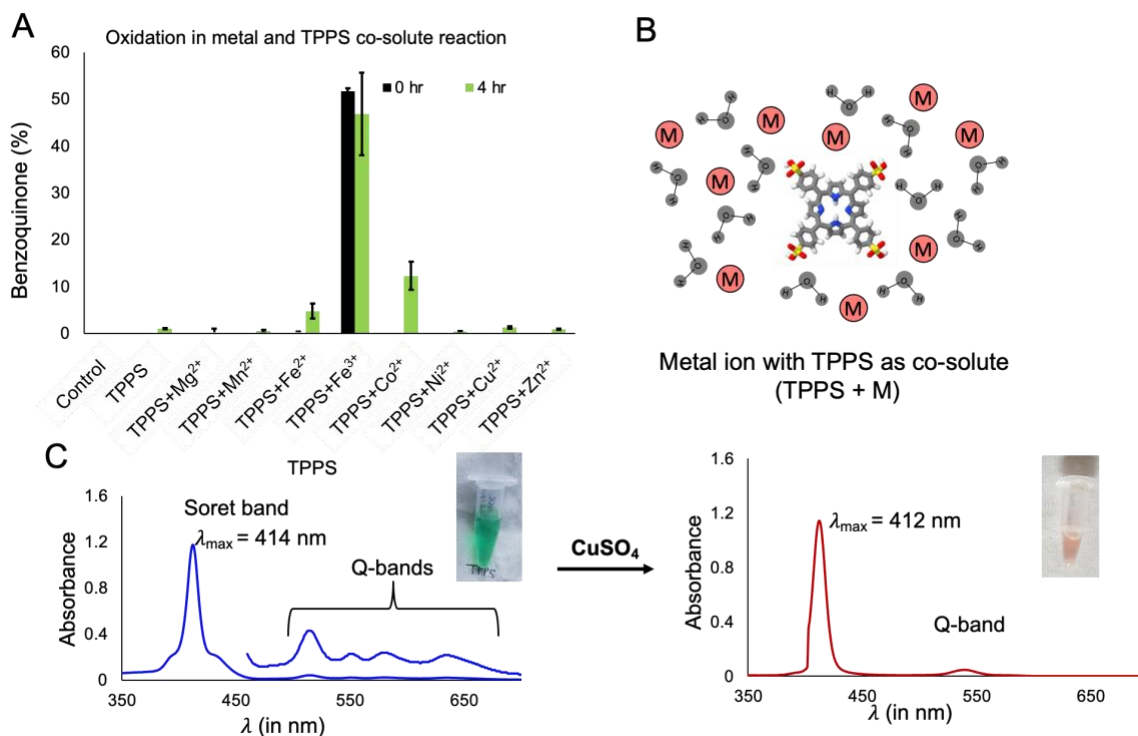
Figure 5.4: Oxidation of HQ to BQ in TPPS and metal ion containing co-solute reactions (panel A). The bar graph shows the percentage of benzoquinone (BQ) produced in the co-solute reactions. X-axis shows the different reactions i.e., only HQ control (control), HQ with TPPS (TPPS), and the different co-solute reactions containing HQ and TPPS along with the different metal ions viz. Mg^{2+} (TPPS+ Mg^{2+}), Mn^{2+} (TPPS+ Mn^{2+}), Fe^{2+} (TPPS+ Fe^{2+}), Fe^{3+} (TPPS+ Fe^{3+}), Co^{2+} (TPPS+ Co^{2+}), Ni^{2+} (TPPS+ Ni^{2+}), Cu^{2+} (TPPS+ Cu^{2+}) and Zn^{2+} (TPPS+ Zn^{2+}), respectively. Error bars depict standard deviation; N=2 (chemical replicates). Black and green colors indicate % BQ produced at the initiation of reaction (0 hr) and at the end (4 hr) of the reaction, respectively. B) Illustration depicting the presence of metal ions and TPPS in the co-solute reaction mixture (M+TPPS). C) Representative UV spectrum of TPPS alone (left panel) and Cu^{2+}-TPPS (right panel), showing changes in the absorbance spectrum on the metal's coordination to porphyrin. Typically, a change in the λ_{max} of the soret band and decrease in the number of Q bands was observed upon complexation.

5.3.3. Prebiotic synthesis of preformed metal-porphyrin complexes

Based on the above observations, we set out to investigate the ability of metal ions to coordinate with TPPS and discern if metal-coordinated TPPS (M-TPPS) complexes could result under prebiotically plausible terrestrial geothermal conditions. In addition to being widely considered as a niche pertinent to discerning the emergence of life on the early Earth, these prebiotic settings have previously been shown to facilitate the formation of biopolymers relevant to extant biochemistry [50–54]. Towards this, an aqueous solution containing 0.3 mM metal ions and 0.03 mM (ten times lower than metal ion) TPPS was incubated at 70° C, simulating putative early Earth conditions [55] (details in Methods section). Metal coordination with the TPPS scaffold was monitored at specific time points using the temporal change observed in the UV absorption spectrum and steady-state fluorescence spectroscopy. The decrease in the intensity at 414 nm (λ_{max} of soret band for TPPS), along with the appearance of other prominent bands with λ_{max} corresponding to M-TPPS complexes, were monitored as an indication of metal coordination [47–49]. A change in λ_{max} was observed for all metals except for Mg^{2+}. The λ_{max} of TPPS (414 nm) shifted to 467 nm for Mn^{2+}, 425 nm for Co^{2+}, 410 nm for Ni^{2+}, 412 nm for Cu^{2+} and 421 nm for Zn^{2+}. Additionally, a decrease in the number of Q-bands was also observed upon complexation. However, Fe^{3+} ions gave rise to two strong bands at 432 nm and 493 nm (Figure 5.11 right panel), with no Q bands in between 500-700 nm when incubated with TPPS. In the case of Mg^{2+}, no change was observed in the absorbance even after thirty-six hours of incubation at 70°C, indicating the absence of coordination.

Upon excitation at 414 nm, uncomplexed TPPS shows emission bands at 640 nm and 697-698 nm [48,56]. However, after coordinating with incompletely filled d-orbital containing metals, its fluorescence gets quenched [48]. This could be due to the paramagnetic nature of these metal ions, which allows excited electrons to lose their energy by intersystem crossing, thereby quenching the fluorescence of TPPS. Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+} and Cu^{2+} ions quenched the fluorescence of TPPS upon coordination due to their paramagnetic nature (Figure 5.5A). This allowed us to examine the M-

TPPS complex formation by monitoring the disappearance of the fluorescence of TPPS for all the metals that we investigated (except Zn^{2+}). In the case of Zn^{2+}, the Zn^{2+}-TPPS complex showed a blue shift in its emission bands (605 nm and 656 nm), upon excitation at 414 nm [48,56]. Therefore, the formation of the Zn^{2+}-TPPS complex was monitored by observing the blue shift of the fluorescence emission bands (Figure 5.5A). Though, in the case of Mg^{2+}, the fluorescence spectrum of TPPS was found to be unaffected even after thirty-six hours of incubation. Further, the formation of coordinated M-TPPS was also confirmed using HRMS (Table 5.2). Our results showed that all the metal ions investigated here, except for Mg^{2+} and Fe^{3+}, could coordinate with TPPS to form M-TPPS under prebiotically pertinent conditions. This observation of Mg^{2+} based reaction was in agreement with what has been reported in a previous study, where it failed to coordinate with uroporphyrin in neutral aqueous conditions [22]. Importantly, the M-TPPS formation even in the micromolar range and under high temperature and aqueous conditions underscores the potential prebiotic plausibility of these M-TPPS complexes.

Table 5.2: High-Resolution Mass spectrometry (HRMS) analysis of formed M-TPPS complexes.

Chemical species	Molecular formula [M]	Species	Calculated Mass	Observed Mass	ppm error
Mn^{2+}-TPPS	$C_{44}H_{28}MnN_4O_{12}S_4$	$[M-4H]^{4-}$	245.7419	245.7441	-8.95
		$[M-4H]^{3-}$	327.9916	327.9947	-9.45
Fe^{3+}-TPPS.OH	$C_{44}H_{29}FeN_4O_{13}S_4$	$[M-4H]^{4-}$	250.4926	250.4938	-4.79
Co^{2+}-TPPS	$C_{44}H_{28}MnN_4O_{12}S_4$	$[M-4H]^{4-}$	246.7407	246.7405	0.811
		$[M-4H]^{3-}$	329.3233	329.3234	-0.30
Ni^{2+}-TPPS	$C_{44}H_{28}MnN_4O_{12}S_4$	$[M-4H]^{4-}$	246.4912	246.4905	2.84
Cu^{2+}-TPPS	$C_{44}H_{28}MnN_4O_{12}S_4$	$[M-4H]^{4-}$	247.7398	247.7386	4.84
Zn^{2+}-TPPS	$C_{44}H_{28}MnN_4O_{12}S_4$	$[M-4H]^{4-}$	247.9897	247.9891	2.42

To investigate the formation kinetics of M-TPPS complexes, fluorescence quenching of TPPS (or shifting in the emission maxima in the case of Zn^{2+}) was monitored (Figure 5.5). TPPS fluorescence was quenched within one hour for Cu^{2+}, Mn^{2+}, Co^{2+}

and Zn^{2+} containing reactions. Peculiarly, the fluorescence intensity was found to increase with an increasing period of incubation in the case of Fe^{3+} ions, confirming the absence of the coordination between Fe^{3+} ions and TPPS (Figure 5.5A). In the case of Ni^{2+}, the TPPS fluorescence was quenched completely only after thirty-six hours of incubation, confirming the formation of Ni-TPPS. This also indicated a lower rate of M-TPPS complex formation in Ni^{2+} containing reactions, when compared to reactions containing Cu^{2+}, Mn^{2+}, Co^{2+} and Zn^{2+} ions. In the case of Cu^{2+}, the fluorescence was quenched immediately after the addition of Cu^{2+} to the TPPS solution (Figure 5.5A). In order to compare the rate of M-TPPS complex formation when using Mn^{2+}, Co^{2+} and Zn^{2+} ions, the fluorescence intensities were monitored for 60 minutes (Figure 5.5 panels B, C and D). The formation of Mn-TPPS was found to reach completion after 30 minutes, whereas for Co^{2+} and Zn^{2+} ions, the coordination was completed within just 15 minutes. Next, to get a more detailed understanding of the M-TPPS formation of the faster-coordinating ions i.e., Cu^{2+}, Zn^{2+} and Co^{2+}, these reactions were monitored continuously at a time interval of 0.1 seconds, using steady-state kinetics (Figure 5.5E). The fluorescence of TPPS was monitored for the first 300 seconds, following which metal salt solution was added to reach a final concentration of 0.3 mM metal ions and 0.03 mM TPPS, and the solution was mixed thoroughly (Figure 5.5E; arrow depicts the addition of metal salt solution). The fluorescence was then monitored for 1500 seconds. The kinetic (fluorescence) decay data was fitted with a first-order exponential decay curve to calculate the decay time constants. For Cu^{2+}, Zn^{2+} and Co^{2+}, the decay time constants were found to be < 4 seconds, 6.56 ± 1.13 seconds and 334.54 ± 42.93 seconds, respectively. The coordination rate of metal ions with TPPS was thus found to be in the following order; Cu^{2+} > Zn^{2+} > Co^{2+} > Mn^{2+} >>> Ni^{2+}. This trend remained the same even in the competition experiments, which contained a mixture of multiple metal ions in a single pot (Figure 5.6).
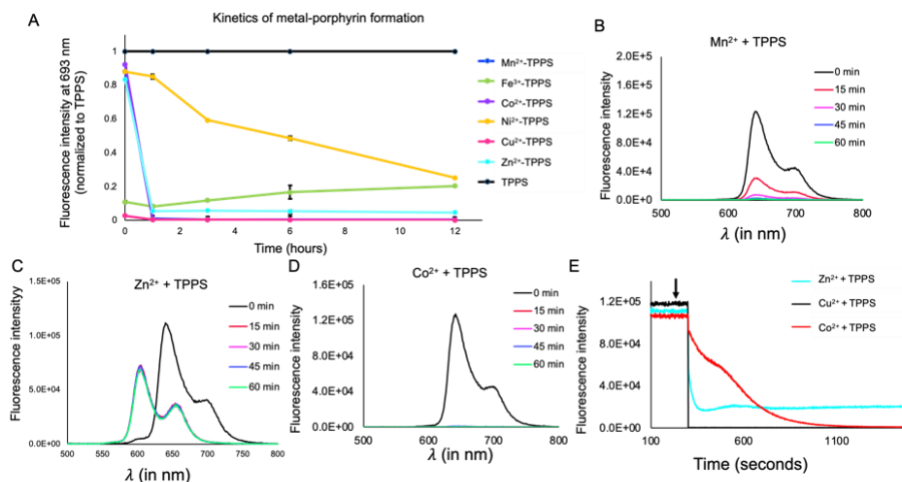
Figure 5.5: Formation kinetics of metal-coordinated TPPS (M-TPPS) complexes using fluorescence spectrum (λ_{ex} = 414 nm). A) Time-based fluorescence quenching of TPPS at 693 nm in the presence of different metal ions. Panels B, C and D show the fluorescence spectrum for the TPPS solution containing Mn^{2+}, Co^{2+} and Zn^{2+}, respectively, after different time periods of incubation (depicted in different colors as indicated above). E) Shows the fast-kinetics for the fluorescence quenching of TPPS at 693 nm in the case of Zn^{2+} (cyan curve), Cu^{2+} (black curve) and Co^{2+} (red curve). Black arrow indicates the timepoint of addition of corresponding metal ions to the TPPS solution after 300 seconds. N = 3 (chemical replicates); Error bars depict standard deviation.
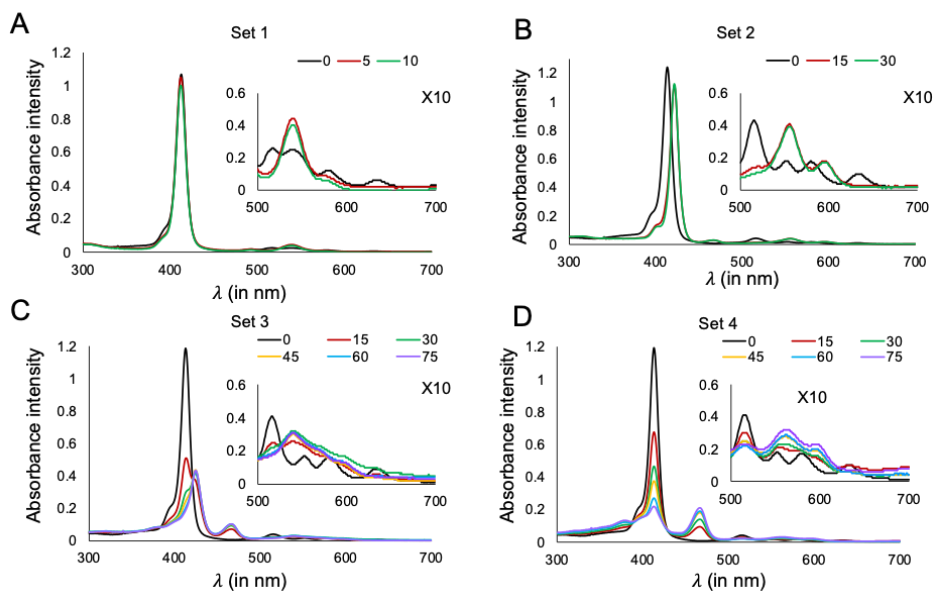
Figure 5.6: Competition experiments for the affinity of different metal ions towards TPPS, monitored using UV spectroscopy at varying time periods as depicted by different colors (in minutes). Different sets indicate mixtures of different metal ions. Legend indicates varying time periods (in minutes). A) **Set 1** shows the preferential formation of Cu^{2+}-TPPS (λ_{max} = 412 nm) when TPPS was incubated with a mixture of **Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}**. B) **Set 2** shows the preferential formation of Zn^{2+}-TPPS (λ_{max} = 421 nm) when TPPS was incubated with a mixture of **Mn^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+} and Zn^{2+}**(Set 1 without Cu^{2+}). C) **Set 3** shows the preferential formation of Co^{2+}-TPPS (λ_{max} = 427 nm) when TPPS was incubated with a mixture of **Mn^{2+}, Fe^{3+}, Co^{2+} and Ni^{2+}** (Set 2 without Zn^{2+}). D) **Set 4** shows the preferential formation of Mn^{2+}-TPPS (λ_{max} = 467 nm) when TPPS was incubated with a mixture of **Mn^{2+}, Fe^{3+} and Ni^{2+}** (Set 3 without Co^{2+}). Inset in all these spectrum shows the 10X zoomed spectra in the region of 500-700 nm (region of Q-bands).

5.3.4. Oxidizing capability of preformed metal complexes

We observed the formation of M-TPPS complexes for all the metal ions investigated except for Mg^{2+} and Fe^{3+} under our reaction conditions. However, the rate of coordination with TPPS varied greatly for different metal ions, ranging from seconds ($Cu^{2+}$and Zn^{2+}) to minutes (Co^{2+} and Mn^{2+}), to even days (Ni^{2+}). Nonetheless, this illustrates the probable availability of the different M-TPPS complexes on the early Earth, as also has been suggested in previous literature [28,29]. Given this, we investigated the oxidizing capability of pre-formed metal-TPPS (M-TPPS) complexes, to characterize the influence of TPPS coordination on the oxidizing activity of the metal ions. Fe^{3+}-TPPS was the only one that was acquired commercially as it failed to form under our reaction conditions.

5.3.4.1. Oxidation of HQ to BQ by different M-TPPS complexes:

No significant HQ oxidation was observed in the presence of Zn^{2+}-TPPS, Mn^{2+}-TPPS and Ni^{2+}-TPPS (Figure 5.7A). For Cu^{2+}-TPPS, up to ~4% of HQ oxidation was

observed after four hours. Interestingly, we observed significantly enhanced HQ oxidation, of up to 46%, in the presence of Co^{2+}-TPPS complex after four hours, showing >4 turnover number (TON, see Methods for more details) (Figure 5.7, panels A and E). This showed that upon coordination with TPPS, the catalytic activity of Co^{2+} ions increased remarkably. In the case of Fe^{3+}-TPPS, ~10.4% HQ oxidation (TON =1) was observed immediately after the addition of the M-TPPS complex, which remained unchanged even after four hours of incubation (Figure 5.7A). However, this percentage of HQ oxidation by Fe^{3+}-TPPS (10.4%) was lower when compared to only (free) Fe^{3+} ion mediated HQ oxidation (~ 30%), and in the co-solute reaction of Fe^{3+} and TPPS (51%) (Figure 5.7E).

The decrease in the activity of Cu^{2+}-TPPS and Fe^{3+}-TPPS when compared to the free metal ions could be because of the formation of near-planar M-TPPS coordinated complexes as was observed in the 3D model. These models were predicted by energy minimization simulations using MolView (Figure 5.8) [57,58]. The presence of the rigid planar macrocycle ring of TPPS is known to restrict the metal's ability to change its stereochemistry and hence, its oxidation state [59]. However, in the case of Co^{2+}, a distortion in the planarity of porphyrin was observed upon coordination in the 3D model, and this can allow the change in the stereochemistry of Co^{2+}-TPPS, which, in turn, can potentially enhance its catalytic ability (Figure 5.8 panels A and B). Nevertheless, the lower oxidizing activity of the coordinated Fe^{3+}-TPPS still does not justify the high oxidation capability that was seen in the co-solute reaction that contained Fe^{3+} and TPPS. Moreover, from UV and fluorescence spectroscopy, it was confirmed that Fe^{3+} was not able to coordinate with TPPS to form the Fe^{3+}-TPPS complex, under our reaction conditions (Figure 5.9 panel A).

To unmask the reason behind the increased catalytic activity of the co-solute reaction containing Fe^{3+} and TPPS, we chose to systematically explore it further. We assessed the possibility of the formation of non-coordinated Fe and TPPS complexes (TPPS**Fe^{3+}) via Fe^{3+} ion interaction with the TPPS moiety. Visible precipitates were also observed in this co-solute reaction upon centrifugation at 5000g for 2 minutes, indicating the presence of large aggregates in this reaction mixture. The formation of

these aggregates was further investigated by looking at the scattering of this solution at 600 nm (the mixture does not absorb light at this wavelength). We observed a drastic increase in the scattering of the TPPS containing solution upon the addition of the Fe^{3+} ions, signifying the formation of higher-order structures (Figure 5.9B). Additionally, the presence of aggregates in the solution was confirmed using optical microscopy (for details see Methods section). Large clusters of rod-shaped aggregates were observed for the mixture containing TPPS and Fe^{3+}, while the control solutions containing Fe^{3+} alone, TPPS alone or preformed Fe^{3+}-TPPS complex by itself (control experiments), failed to show any detectable aggregation (Figure 5.10). From previous literature, it is known that porphyrin complexes predominantly form side-by-side aggregates (J-aggregates) and plane-to-plane aggregates (H-aggregates) [56,60–63]. Studies have also suggested that the catalytic activity of porphyrin increases upon aggregation [61,62,64]. Given this, it seemed like a reasonable premise to base our assumption that the increase in the catalytic activity in the co-solute reactions containing Fe^{3+} and TPPS, could possibly be because of the formation of such aforementioned aggregates.
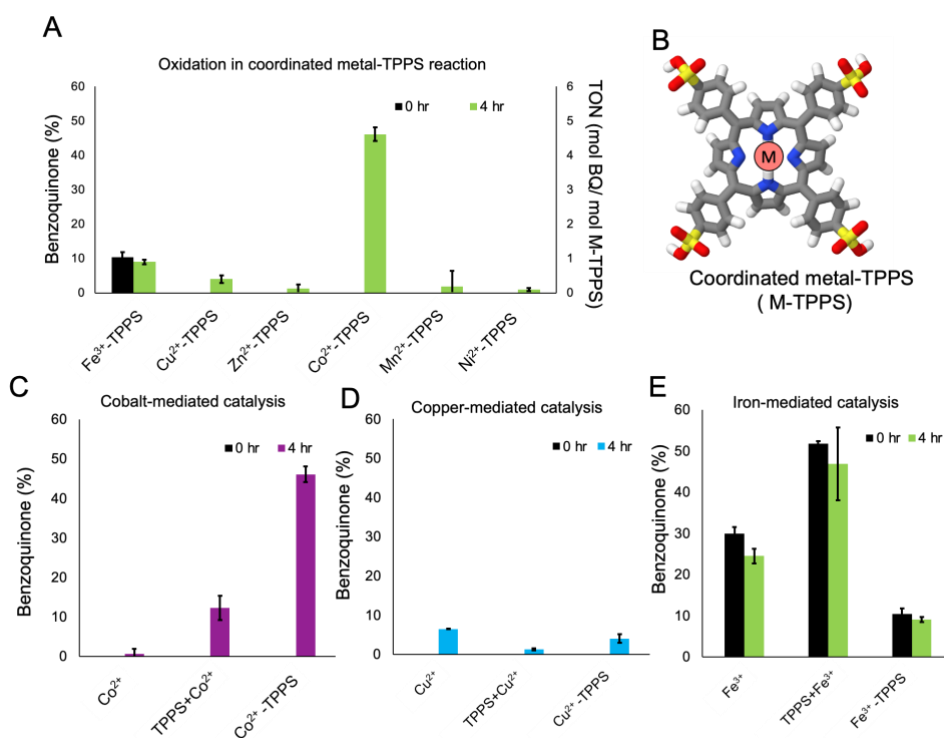
Figure 5.7: Oxidation of HQ by coordinated metal-TPPS (M-TPPS) complexes. A) Reactions of HQ with the respective M-TPPS complexes. Left Y-axis shows BQ (%) produced and the right Y-axis shows the turnover number (TON) at the initiation (0 hr, black bars) and at the end of the reaction (4 hr, green bars). B) Illustration depicting coordinated metal -TPPS (M-TPPS) complex. Panels C-E show the comparison of oxidation by free metal ions vs the respective metal ion with TPPS in a co-solute reaction vs the coordinated porphyrin complex (indicated in the X-axis). Time points shown are from the initiation of the reaction (black bars) and after four hours of incubation for Co^{2+} (panel B; purple bars), Cu^{2+} (panel C; blue bars) and Fe^{3+} (panel D; green bars) based reactions, respectively. N=3 (chemical replicates); Error bars depict standard deviation.
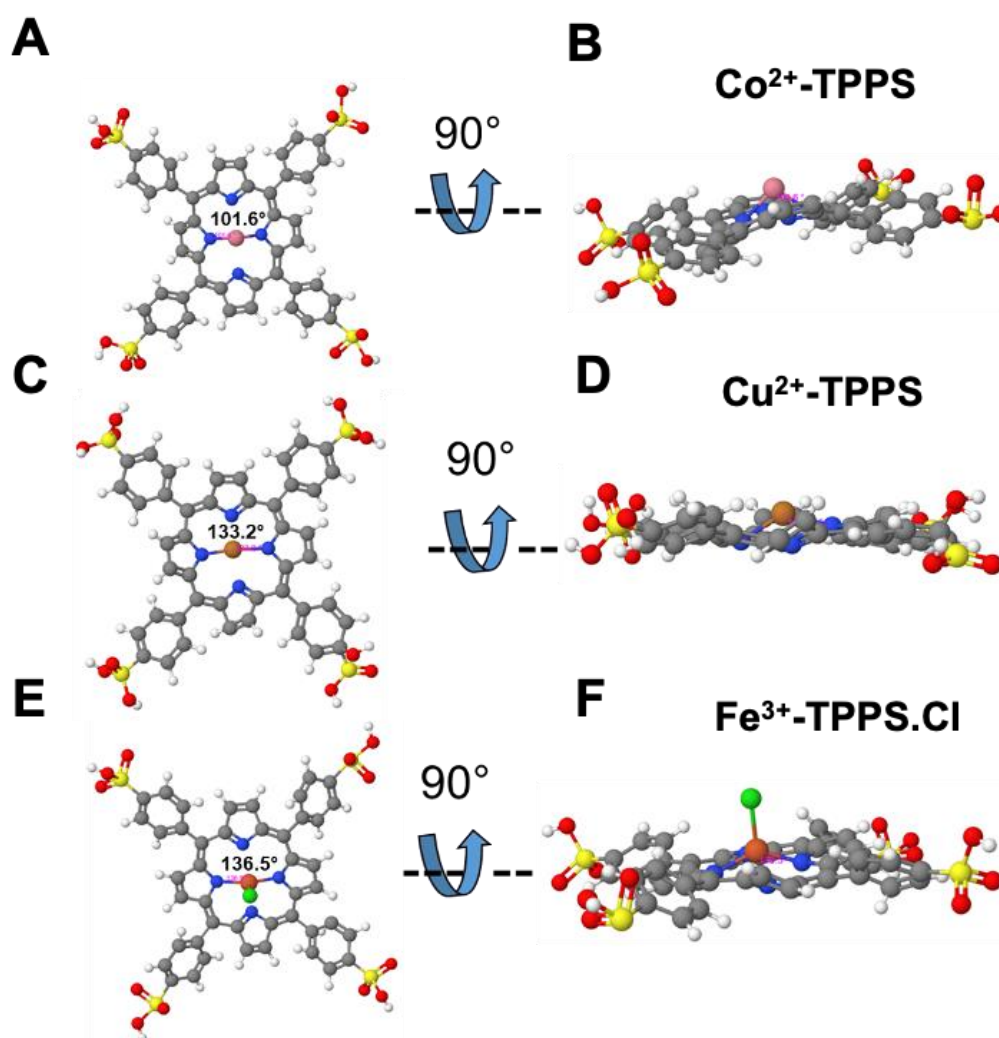
Figure 5.8: Predicted 3D model of M-TPPS (M=Co^{2+}, Cu^{2+} and Fe^{3+}) complex using energy-minimization simulations in MolView (molview.org). A, C and E represent Co^{2+}-TPPS, Cu^{2+}-TPPS and Fe^{3+}-TPPS.Cl, respectively; B, D and F represent 90° rotated version of A, C and E, respectively. The angle between the N-M-N in the core for Co^{2+}, Cu^{2+} and Fe^{3+} was 101.6°, 133.2° and 136.5°.

Upon formation of J-aggregates and H-aggregates, the soret absorption band of porphyrin is known to undergo red-shift and blue shift, respectively [60,61,65,66]. The absorption λ_{max} of J-aggregates of TPPS was reported to be near ~ 490 nm in aqueous solution [60,61,65,66]. The co-solute reaction containing Fe^{3+} and TPPS showed a red-shifted band with a λ_{max} at 494 nm, alluding to the formation of TPPS**Fe^{3+} J-aggregates (Figure 5.9A). The presence of J-aggregates was further confirmed by Field emission Scanning Electron Microscopy (FESEM), followed by the elemental analysis of the observed aggregates (Figure 5.9 panels C and D). The presence of elongated planar structures confirmed the presence of J-aggregates [61]. Further, the elemental analysis of these structures showed the presence of Fe, C, N, O and S, indicating that the structures visible in the viewing field were indeed aggregates containing Fe^{3+} and TPPS moieties. Moreover, prolonged heating of the TPPS**Fe^{3+} aggregates at 100°C, resulted in the diminishing of the absorption peak at 494 nm that corresponds to the J-aggregates (Figure 5.10). This could result from the instability of the aggregated species at this high temperature, resulting in the formation of Fe^{3+}-TPPS coordinated complex instead. Given this, in the case of Fe^{3+}-TPPS formation, it seems that the coordination could result via the attack of Fe^{3+} ion on the aggregated TPPS at elevated temperature (100°C). Interestingly, such a mechanism has been proposed for the formation of coordinated Cu^{2+}-porphyrin complex in the presence of porphyrins in its aggregated form (J-aggregates) [67]. Additionally, whether the other metal ions studied here (Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}) could also form similar aggregates with TPPS, was evaluated by looking at the scattering of their respective solutions (Figure 5.9B). No increase in solution scattering, along with absence of visual aggregates upon centrifugation, and the

absence of absorption peaks corresponding to J- and H-aggregates, eliminated the possibility of the other metal ions forming non-covalent aggregates with TPPS.
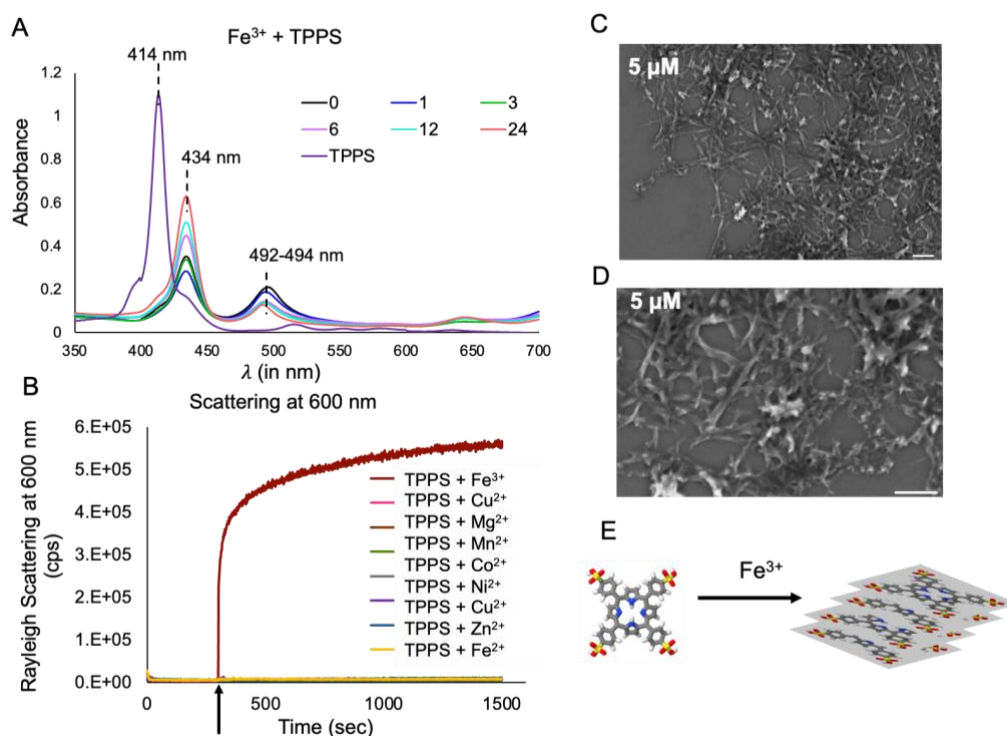

Figure 5.9: Characterization of Fe^{3+} and TPPS non-coordinated aggregates (TPPS**Fe^{3+}). A) UV absorbance spectrum of co-solute reaction containing Fe^{3+} and TPPS, after different time periods (at the initiation (0), after 1 hour (1), 3 hours (3), 6 hours (6), 12 hours (12) and 24 hours (24), respectively) along with TPPS as reference. Different colors depict different time periods as shown in the legend; purple color is the UV spectrum of the reference TPPS sample. The red shift of 414 nm soret band to 434 nm and 492 nm bands was observed, indicating the formation of J-aggregates. B) Rayleigh scattering at 600 nm confirming the formation of higher order structures only in the Fe^{3+} and TPPS containing co-solute reaction. The black arrow indicates the time when the addition of different metal ions to the TPPS solution was done. Different colors represent the kinetic curve obtained after the addition of the different metal ions (as shown in the Figure legend). All the readings were repeated three times, (N=3, chemical replicates). Panel C shows the Field Emission Scanning electron microscopy (FESEM) image of 5 µM of TPPS solution containing 50 µM of Fe^{3+} ions, along with its magnified

version in panel D. The scale bar is 500 nm. The appearance of elongated rod-like structures confirmed the presence of J-aggregates. E) Illustration showing the Fe^{3+} induced formation of J-aggregates of TPPS.
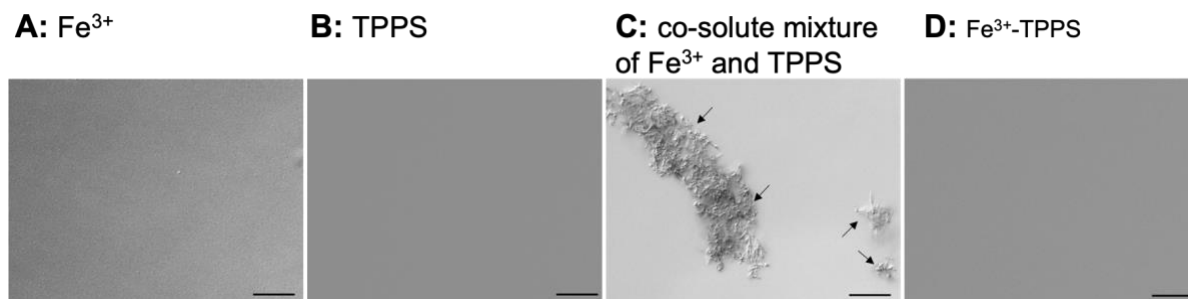

Figure 5.10: Differential Interference Contrast (DIC) microscopy (under 40X objective). A) 300 µM of Fe^{3+} solution. B) 30 µM TPPS solution. C) co-solute mixture of 300 µM Fe^{3+} and 30 µM TPPS. D) 30 µM Fe^{3+}-TPPS solution. Black arrows indicate towards the non-coordinated aggregates. Scale bar = 20 µm.
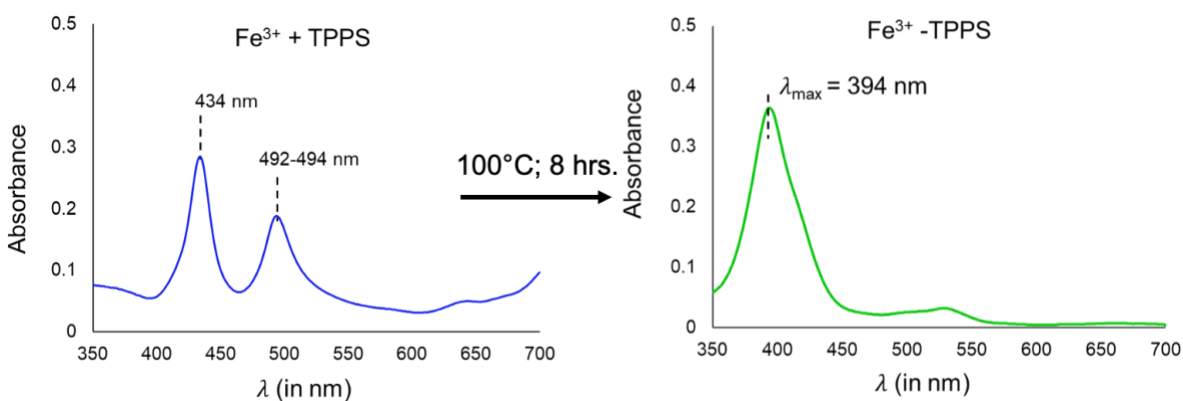
Figure 5.11: UV spectrum of the preformed Fe^{3+}-TPPS coordinated complex with λ_{max} at 394 nm (green spectrum, right panel) that results upon heating of the co-solute mixture of 0.3 mM Fe^{3+} and 0.03 mM TPPS (blue spectrum, left panel), at 100ºC for 8 hours.

5.3.4.2. Oxidation of NADH to NAD by M-TPPS complexes:

The spontaneous emergence of catalysts on the early Earth would have been governed by a multitude of factors, making their ready emergence a non-trivial

exercise [3]. To circumvent this, early protoenzymes could have possibly evolved the ability to catalyze a wider array of prebiotically relevant reactions. In this context, we investigated the ability of M-TPPS to catalyze the oxidation of NADH at 25° C (see Methods section). Along with quinones, NADH is another prebiotically plausible cofactor that is involved in extant redox reactions and is a key molecule in metabolism [15,17,68,69]. This was also chosen as a representative for nucleotide-based cofactors like flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN), which are utilized extensively in contemporary biology. NAD has a characteristic UV absorption band at 260 nm. On the other hand, NADH absorbs at 340 nm in addition to 260 nm. Therefore, in order to monitor the oxidation of NADH, the disappearance of 340 nm absorption signal was monitored. In the control reactions, i.e., in the absence of M-TPPS, only 1.5% oxidation was observed even after four hours of incubation, indicating negligible spontaneous oxidation of NADH. Nonetheless, our M-TPPS experiments showed that up to 12.3% NADH was oxidized immediately after the addition of Co^{2+}-TPPS to the NADH solution (Figure 5.12A). This NADH oxidation was observed to increase gradually and reach up to 59.9% after four hours. Additionally, the control reactions containing either the free Co^{2+} ions or just the TPPS scaffold showed negligible oxidation. Along with Co^{2+}-TPPS, other M-TPPS were also observed to oxidize NADH to varying extent. The oxidation ability of different M-TPPS towards NADH followed this order: Co^{2+}-TPPS (59.9%) > Cu^{2+}-TPPS (52.8%) > Fe^{3+}-TPPS (50.4%) > Mn^{2+}-TPPS (29.3%) > Zn^{2+}-TPPS (26.9%) > Ni^{2+}-TPPS (12.9%) (Figure 5.12). Also, the M-TPPS mediated NADH oxidation displayed different TON ranging from ~6 for Co^{2+}-TPPS to >1 for Ni^{2+}-TPPS (Figure 5.12). The capability of M-TPPS to also oxidize NADH, emphasizes their versatility for using multiple substrates; a property that would have been hugely advantageous on the prebiotic Earth. This is pertinent as it would have taken a protracted process of molecular evolution, over very long-time scales, that would have subsequently allowed for the emergence of highly specific and efficient catalysts for individual substrates.
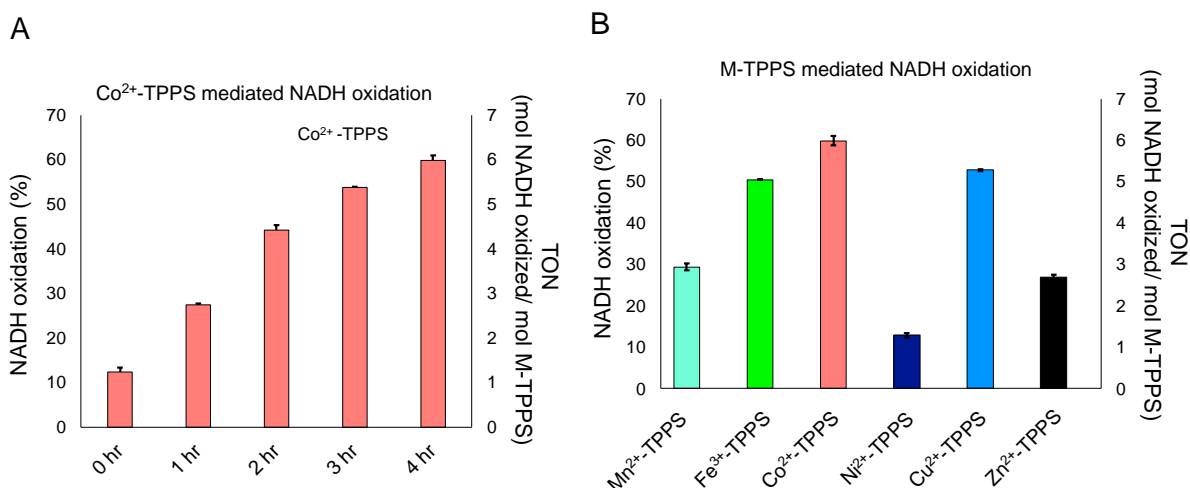
Figure 5.12: Oxidation of NADH mediated by preformed M-TPPS complexes. A) Co^{2+}-TPPS mediated NADH oxidation after different time periods (salmon bars). X-axis shows the different time periods i.e., at the initiation of the reaction (0 hr), after one hour (1 hr), two hours (2 hr), three hours (3 hr) and four hours (4 hr) of incubation, respectively, at 25°C. Left Y-axis depicts the percentage of NADH oxidized and right Y-axis depicts the turn over number (TON) of Co^{2+}-TPPS. B) Comparison of NADH oxidation mediated by different M-TPPS. Left Y-axis depicts the percentage of NADH oxidized and right Y-axis depicts the turn over number (TON) of the M-TPPS. N=2 (chemical replicates); error bars depict standard deviation.

5.4. Discussion

Porphyrin-containing metalloenzymes are hypothesized to be some of the most ancient catalysts that are thought to have been present at the dawn of life on the early Earth [1,16,70]. However, their catalytic ability in prebiotically pertinent reactions, has largely remained unexplored. In this work, we investigated the influence of porphyrin scaffolds, on the modulation of oxidation ability of different metal ions. In the case of free metal ion-mediated reactions, the highest oxidation was observed for Fe^{3+} followed by Cu^{2+}, Fe^{2+} and Mn^{2+} ions (Figure 5.3). In the co-solute reactions, the HQ oxidation was found to be increased for Fe^{3+} (from 30% to 51.7%) and Co^{2+} (from negligible to 12.3%) ions in the presence of TPPS, when compared to just the

corresponding free metal ion-mediated oxidation (Figure 5.7). Contrarily, the HQ oxidation by Cu^{2+} was found to diminish in the presence of TPPS while it remained unaltered for the rest of the metal ions. Given this data, we conclude that the presence of TPPS as a co-solute increase the catalyzing activity of Fe^{3+} and Co^{2+} ions, while damping the activity of Cu^{2+} ions.
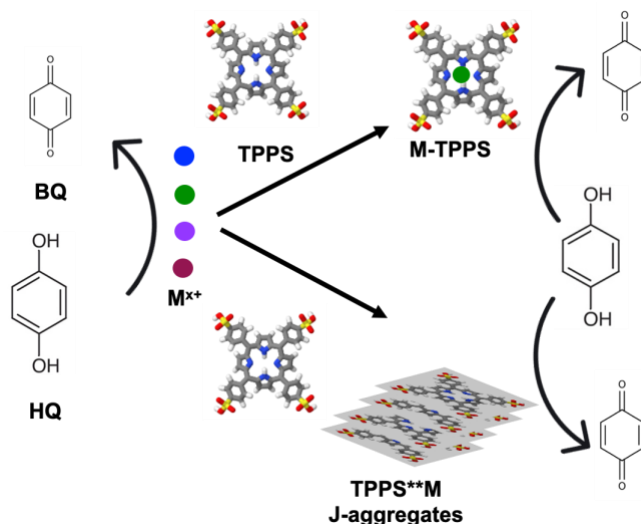

Figure 5.13: Schematic showing the two different pathways described in this study by which oxidation of HQ to BQ is made feasible. Different color spheres show the different metal ions. These metal ions in the presence of TPPS can then either coordinate to form M-TPPS complex or induce J-aggregate formation (TPPS**M; a non-coordinated complex). These complexes (coordinated or non-coordinated) have different catalytic efficiencies towards HQ oxidation.

Changes in the UV-absorption and fluorescence emission profiles in the co-solute reactions, indicated that the change in the catalytic activity of the metal ions in the presence of TPPS might possibly stem from the interaction between the metal ions and the TPPS. This was further confirmed by similar changes that were seen in the activity of the pre-formed coordinated complexes of Cu^{2+}-TPPS and Co^{2+}-TPPS mediated oxidation reactions. In contrast to Co^{2+}, the inactivity of Cu^{2+}-TPPS could be attributed to the inability of copper to change its oxidation state when it is coordinated to TPPS, which is a planar rigid structure [59]. Importantly, this observation highlighted that the catalytic activity of metal ions, does not necessarily increase upon

complexation, but rather can vary greatly depending on the metal ions. This underscores the interesting possibility that the coordination/interaction of metals with organic scaffolds such as TPPS could have indeed acted as a selectivity barrier for the formation of metal-organic scaffolds. Importantly, the preformed M-TPPS complexes were also found to catalyze the oxidation of NADH to NAD, signifying its substrate versatility. Similar to HQ based reactions, the highest oxidation of NADH was observed for Co^{2+}-TPPS. This capability to recognize different types of substrates would have been beneficial for primitive metalloenzymes.

While characterizing the prebiotic plausibility of metalloporphyrins, we observed that Cu^{2+}, Co^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+} ions were able to coordinate with TPPS under our laboratory simulated prebiotic settings. However, the rate of the formation of M-TPPS varied greatly as coordination with TPPS could occur in a few seconds (as was observed for Cu^{2+}, Zn^{2+} and Co^{2+}), or could even happen on the order of days (as was seen for Ni^{2+}). Therefore, the rate of formation of such metal coordinated complexes could have also acted as another important checkpoint, wherein the selection of certain metal ions to form metalloporphyrins could have happened.

Interestingly, the oxidizing capability of the Fe^{3+}-TPPS (coordinated) complex was found to be lower when compared to HQ oxidation that occurred in the presence of free Fe^{3+} ions or the co-solute reactions that contained Fe^{3+} ions and TPPS. This is because, instead of coordinating with TPPS, Fe^{3+} was found to form non-coordinated aggregates with TPPS under our reaction conditions. From UV spectroscopy, Rayleigh scattering and FESEM, Fe^{3+} ions were shown to assemble into J-aggregates with TPPS, which potentially resulted in an enhanced oxidation activity (increased by 1.7-fold when compared to the free Fe^{3+} ion containing reaction). Interestingly, ~ 27% of extant oxidoreductase enzymes utilize Fe^{3+}/ Fe^{2+} redox couple to catalyze reactions [71,72]. Also, previous studies have indicated that the metal ion availability in the Archean era would have followed the order mentioned here: $Fe^{+3/+2}$> Mn^{+2}, Ni^{+2}, Co^{+2}>> Cd^{+2}, Zn^{+2}> Cu^{+2} [73,74]. Given this abundance of $Fe^{+3/+2}$ on the prebiotic Earth, along with its ability to catalyze reactions on its own, it

could have readily induced aggregation of porphyrin like scaffolds, resulting in non-coordinated catalytically active metal-organic scaffolds as we show in this study.

To summarize, our work demonstrates how a change in the oxidizing activity of different metal ions can be brought about by coordination/interaction with prebiotically plausible inorganic scaffolds (e.g. TPPS) (Figure 5.13). We showed the influence of porphyrin in modulating the oxidizing ability of free metal ions, either in a positive or negative manner, depending on the metal ion in question. We also demonstrated the formation of different M-TPPS (coordinated) complexes and Fe^{3+} containing J-aggregates under prebiotically pertinent conditions. The eventual refinement of the usage of a specific ion type, in conjunction with a suitable scaffold, would have resulted in highly precise and effective catalysts as seen in contemporary biology. This refinement would have potentially resulted over millions/billions of years of evolution. Various factors, starting from the availability of the metal ions, formation of metal-organic scaffolds, catalytic activity and physicochemical environment, would have played an important role in the formation of primitive metalloenzymes. In this context, this study underscores the central role of the complex interactions and cross-talk that would have occurred in a heterogenous prebiotic soup between metal ions and available scaffolds and their potential role in the emergence of early metalloporphyrins.

5.5. References

1. Raanan, H., Pike, D. H., Moore, E. K., Falkowski, P. G. & Nanda, V. Modular origins of biological electron transfer chains. *Proc. Natl. Acad. Sci.* **115**, 1280–1285 (2018).

2. Holm, R. H., Kennepohl, P. & Solomon, E. I. Structural and functional aspects of metal sites in biology. *Chem. Rev.* **96**, 2239–2314 (1996).

3. Muchowska, K. B., Varma, S. J. & Moran, J. Nonenzymatic Metabolic Reactions and Life's Origins. *Chem. Rev.* **120**, 7708–7744 (2020).

4. Kaddour, H. *et al.* Nonenzymatic RNA Oligomerization at the Mineral-Water Interface: An Insight into the Adsorption-Polymerization Relationship. *J. Phys. Chem. C* **122**, 29386–29397 (2018).

5. Wächtershäuser, G. Groundworks for an evolutionary biochemistry: The iron-sulphur world. *Prog. Biophys. Mol. Biol.* **58**, 85–201 (1992).

6. Waldron, K. J. & Robinson, N. J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol.* **7**, 25–35 (2009).

7. Muchowska, K. B. *et al.* Metals promote sequences of the reverse Krebs cycle. *Nat. Ecol. Evol.* **1**, 1716–1721 (2017).

8. Keller, M. A., Turchyn, A. V. & Ralser, M. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. *Mol. Syst. Biol.* **10**, 1–12 (2014).

9. Coggins, A. J. & Powner, M. W. Prebiotic synthesis of phosphoenol pyruvate by α-phosphorylation-controlled triose glycolysis. *Nat. Chem.* **2**, 1–76 (2016).

10. Keller, M. A. *et al.* Conditional iron and pH-dependent activity of a non-enzymatic glycolysis and pentose phosphate pathway. *Sci. Adv.* **2**, e1501235 (2016).

11. Hanczyc, M. M., Mansy, S. S. & Szostak, J. W. Mineral surface directed membrane assembly. *Orig. Life Evol. Biosph.* **37**, 67–82 (2007).

12. Siegel, B. Z. & Siegel, S. M. Enzyme-Mimicking properties of silicates and other minerals. *Adv. Sp. Res.* **1**, 27–36 (1981).

13. Mateo-Marti, E., Galvez-Martinez, S., Gil-Lozano, C. & Zorzano, M.-P. Pyrite-induced uv-photocatalytic abiotic nitrogen fixation: implications for early

atmospheres and Life. *Sci. Rep.* **9**, 15311 (2019).

14.   Sanchez-Arenillas, M. & Mateo-Marti, E. Pyrite surface environment drives molecular adsorption: cystine on pyrite(100) investigated by X-ray photoemission spectroscopy and low energy electron diffraction. *Phys. Chem. Chem. Phys.* **18**, 27219–27225 (2016).

15.   Bonfio, C. *et al.* Prebiotic iron–sulfur peptide catalysts generate a pH gradient across model membranes of late protocells. *Nat. Catal.* **1**, 616–623 (2018).

16.   Raanan, H., Poudel, S., Pike, D. H., Nanda, V. & Falkowski, P. G. Small protein folds at the root of an ancient metabolic network. *Proc. Natl. Acad. Sci.* **117**, 7193–7199 (2020).

17.   Goldman, A. D. & Kacar, B. Cofactors are Remnants of Life's Origin and Early Evolution. *J. Mol. Evol.* **89**, 127–133 (2021).

18.   Gibney, B. R., Mulholland, S. E., Rabanal, F. & Dutton, P. L. Ferredoxin and ferredoxin-heme maquettes. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15041–15046 (1996).

19.   Scintilla, S. *et al.* Duplications of an iron-sulphur tripeptide leads to the formation of a protoferredoxin. *Chem. Commun.* **52**, 13456–13459 (2016).

20.   Kadish, K. M., Smith, K. M. & Guilard, R. *The Porphyrin Handbook. The Porphyrin Handbook* **13**, (2012).

21.   Liu, J. *et al.* Metalloproteins Containing Cytochrome , Iron − Sulfur , or Copper Redox Centers. *Chem. Rev.* **114**, 4366–4469 (2014).

22.   Soares, A. R. M., Anderson, D. R., Chandrashaker, V. & Lindsey, J. S. Catalytic diversification upon metal scavenging in a prebiotic model for formation of tetrapyrrole macrocycles. *New J. Chem.* **37**, 2716–2732 (2013).

23.   Fabbri, C., Aurisicchio, C. & Lanzalunga, O. Iron porphyrins-catalysed oxidation of α-alkyl substituted mono and dimethoxylated benzyl alcohols. *Cent. Eur. J. Chem.* **6**, 145–153 (2008).

24.   Calvin, M. Evolution of Enzymes and the Photosynthetic Apparatus. *Science (80).* **130**, 1170–1174 (1959).

25.   Belmonte, L. & Mansy, S. S. Metal catalysts and the origin of life. *Elements* **12**, 413–418 (2016).

26. Hayatsu, R., Studier, M. H., Matsuoka, S. & Anders, E. Origin of organic matter in early solar system-VI. Catalytic synthesis of nitriles, nitrogen bases and porphyrin-like pigments. *Geochim. Cosmochim. Acta* **36**, 555–571 (1972).

27. Hodgson, G. W. & Ponnamperuma, C. PREBIOTIC PORPHYRIN GENESIS: PORPHYRINS FROM ELECTRIC DISCHARGE IN METHANE, AMMONIA, AND WATER VAPOR. *Proc. Natl. Acad. Sci.* **59**, 22–28 (1968).

28. Fox, S. & Strasdeit, H. A possible prebiotic origin on volcanic islands of oligopyrrole-type photopigments and electron transfer cofactors. *Astrobiology* **13**, 578–595 (2013).

29. Pleyer, H. L., Strasdeit, H. & Fox, S. A Possible Prebiotic Ancestry of Porphyrin-Type Protein Cofactors. *Orig. Life Evol. Biosph.* **48**, 347–371 (2018).

30. Lindsey, J. S., Ptaszek, M. & Taniguchi, M. Simple Formation of an Abiotic Porphyrinogen in Aqueous Solution. *Orig. Life Evol. Biosph.* **39**, 495–515 (2009).

31. Premović, P. I., Pavlović, M. S. & Pavlović, N. Z. Vanadium in ancient sedimentary rocks of marine origin. *Geochim. Cosmochim. Acta* **50**, 1923–1931 (1986).

32. Mauzerall, D. C. Evolution of porphyrins. *Clin. Dermatol.* **16**, 195–201 (1998).

33. Anbar, A. D. & Knoll, A. H. Proterozoic Ocean Chemistry and Evolution: A Bioinorganic Bridge? *Science (80).* **297**, (2002).

34. Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L. & Thornton, J. M. Metal ions in biological catalysis: From enzyme databases to general principles. *J. Biol. Inorg. Chem.* **13**, 1205–1218 (2008).

35. Eichhorn, G. L. Metal Ion Catalysis in Biological Systems. *Adv. Chem.* **37**, 37–55 (1962).

36. Karlin, K. D. Metalloenzymes, structural motifs, and inorganic models. *Science (80).* **261**, 701–708 (1993).

37. Bonfio, C. *et al.* UV-light-driven prebiotic synthesis of iron-sulfur clusters. *Nat. Chem.* **9**, 5907 (2017).

38. Huynh, M. T., Anson, C. W., Cavell, A. C., Stahl, S. S. & Hammes-Schiffer, S. Quinone 1 e- and 2 e-/2 H+ Reduction Potentials: Identification and Analysis of Deviations from Systematic Scaling Relationships. *J. Am. Chem. Soc.* **138**,

15903–15910 (2016).

39. Bard & AllenJ. *Standard Potentials in Aqueous Solution*. (Routledge, 2017).

40. Milazzo, G., Caroli, S. & Braun, R. D. Tables of Standard Electrode Potentials. *J. Electrochem. Soc.* **125**, 261C (1978).

41. Swift, E. H. & Butler, E. A. *Quantitative measurements and chemical equilibria*. (W. H. Freeman, 1972).

42. Varma, S. J., Muchowska, K. B., Chatelain, P. & Moran, J. Native iron reduces CO2 to intermediates and end-products of the acetyl-CoA pathway. *Nat. Ecol. Evol.* **2**, 1019–1024 (2018).

43. Frenkel-Pinter, M., Sargon, A. B., Glass, J. B., Hud, N. V. & Williams, L. D. Transition metals enhance prebiotic depsipeptide oligomerization reactions involving histidine. *RSC Adv.* **11**, 3534–3538 (2021).

44. Kitadai, N. *et al.* Metals likely promoted protometabolism in early ocean alkaline hydrothermal systems. *Sci. Adv.* **5**, eaav7848 (2019).

45. Messner, C. B., Driscoll, P. C., Piedrafita, G., De Volder, M. F. L. & Ralser, M. Nonenzymatic gluconeogenesis-like formation of fructose 1,6-bisphosphate in ice. *Proc. Natl. Acad. Sci.* **114**, 7403–7407 (2017).

46. Damer, B. & Deamer, D. The hot spring hypothesis for an origin of life. *Astrobiology* **20**, 429–452 (2020).

47. Song, Z., Adeyemo, A. O., Baker, J., Traylor, S. M. & Lightfoot, M. L. Structure of Porphyrin Tpps4 and Its Interaction With Metal Ions As Elucidated By 1H Nmr and Uv-Visible Spectra. *Georg. J. Sci.* **69**, 89–101 (2011).

48. Beyene, B. B., Yibeltal, A. W. & Ayana, M. T. Colorimetric and fluorescent on-off detection of Cu2+, Sn2+ and Zn2+ by a water-soluble porphyrin: Electronic absorption and emission study. *Results Chem.* **2**, 100058 (2020).

49. Beyene, B. B., Mane, S. B. & Hung, C.-H. Highly efficient electrocatalytic hydrogen evolution from neutral aqueous solution by a water-soluble anionic cobalt(II) porphyrin. *Chem. Commun* **51**, 15067 (2015).

50. Joshi, M. P., Sawant, A. A. & Rajamani, S. Spontaneous emergence of membrane-forming protoamphiphiles from a lipid–amino acid mixture under wet–dry cycles. *Chem. Sci.* **12**, 2970–2978 (2021).

51. Mungi, C. V., Rajamani, S., Mungi, C. V. & Rajamani, S. Characterization of RNA-Like Oligomers from Lipid-Assisted Nonenzymatic Synthesis: Implications for Origin of Informational Molecules on Early Earth. *Life (Basel, Switzerland)* **5**, 65–84 (2015).

52. Becker, S. *et al.* Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat. Commun.* **9**, 1–9 (2018).

53. Sarkar, S., Dagar, S. & Rajamani, S. Influence of Wet–Dry Cycling on the Self-Assembly and Physicochemical Properties of Model Protocellular Membrane Systems. *ChemSystemsChem* **3**, (2021).

54. Dagar, S., Sarkar, S. S. & Rajamani, S. Geochemical influences on nonenzymatic oligomerization of prebiotically relevant cyclic nucleotides. *RNA* **26**, 756–769 (2020).

55. Deamer, D., Singaram, S., Rajamani, S., Kompanichenko, V. & Guggenheim, S. Self-assembly processes in the prebiotic environment. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361**, 1809–18 (2006).

56. Kalyanasundaram, K. & Neumann-Spallart, M. Photophysical and redox properties of water-soluble porphyrins in aqueous media. *J. Phys. Chem.* **86**, 5163–5169 (1982).

57. Smith, T. J. MOLView: A program for analyzing and displaying atomic structures on the Macintosh personal computer. *J. Mol. Graph.* **13**, 122–125 (1995).

58. Bergwerf, H. *MolView : an attempt to get the cloud into chemistry classrooms*. *ACS CHED CCCE Newsletter* (2015).

59. Kadish, K. M. & Van Caemelbecke, E. Electrochemistry of porphyrins and related macrocycles. *J. Solid State Electrochem.* **7**, 254–258 (2003).

60. Jiang, S., Zhang, L. & Liu, M. Photo-triggered J-aggregation and chiral symmetry breaking of an anionic porphyrin (TPPS) in mixed organic solvent. *Chem. Commun.* 6252–6254 (2009). doi:10.1039/b912825e

61. Devaramani, S. *et al.* Porphyrin aggregates decorated MWCNT film for solar light harvesting: influence of J- and H-aggregation on the charge recombination resistance, photocatalysis, and photoinduced charge transfer kinetics. *Phys. Chem. Chem. Phys.* **19**, 18232–18242 (2017).

62. Charron, D. M. *et al.* Photophysics of J-Aggregating Porphyrin-Lipid Photosensitizers in Liposomes: Impact of Lipid Saturation. *Langmuir* **36**, 5385–5393 (2020).

63. Tokunaga, E. & Nakata, K. Giant electrooptic effect of porphyrin J-aggregates in polymer film and in Aqueous solution. *J-Aggregates* **2**, 213–246 (2012).

64. Wang, S. *et al.* Design and construction of artificial photoresponsive protocells capable of converting day light to chemical energy. *J. Mater. Chem. A* **5**, 24612–24616 (2017).

65. Lu, J., Li, Z., An, W., Liu, L. & Cui, W. Tuning the Supramolecular Structures of Metal-Free Porphyrin via Surfactant Assisted Self-Assembly to Enhance Photocatalytic Performance. *Nanomaterials* **9**, 1321 (2019).

66. Siggel, U. *et al.* Photophyshl and Photochemical Properties of Porphyrin Aggregates. *Berichte der Bunsengesellschaft für Phys. Chemie* **100**, 2070–2075 (1996).

67. Trapani, M. *et al.* Mechanism for Copper(II)-Mediated Disaggregation of a Porphyrin J-Aggregate. *ACS Omega* **3**, 18843–18848 (2018).

68. Cleaves, H. J. & Miller, S. L. The Nicotinamide Biosynthetic Pathway Is a By-Product of the RNA World. *J. Mol. Evol.* **52**, 73–77 (2001).

69. Kim, H.-J. J. & Benner, S. A. A Direct Prebiotic Synthesis of Nicotinamide Nucleotide. *Chem. - A Eur. J.* **24**, 581–584 (2018).

70. Harel, A., Bromberg, Y., Falkowski, P. G. & Bhattacharya, D. Evolutionary history of redox metal-binding domains across the tree of life. *Proc. Natl. Acad. Sci.* **111**, 7042–7047 (2014).

71. Putignano, V., Rosato, A., Banci, L. & Andreini, C. MetalPDB in 2018: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res.* **46**, D459–D464 (2018).

72. Andreini, C., Cavallaro, G., Lorenzini, S. & Rosato, A. MetalPDB: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res.* **41**, D312–D319 (2012).

73. Saito, M. A., Sigman, D. M. & Morel, F. M. M. The bioinorganic chemistry of the ancient ocean: The co-evolution of cyanobacterial metal requirements and

biogeochemical cycles at the Archean-Proterozoic boundary? *Inorganica Chim. Acta* **356**, 308–318 (2003).

74.   Anbar, A. Elements and evolution. *Science (80).* **332**, 1481–1483 (2008).

PUBLICATIONS

1. Dagar, Shikha, Susovan Sarkar, and Sudha Rajamani. "Geochemical influences on nonenzymatic oligomerization of prebiotically relevant cyclic nucleotides." *RNA* 26.6 (2020): 756-769.

2. Dagar, Shikha, Susovan Sarkar, and Sudha Rajamani. "Porphyrin in prebiotic catalysis: Ascertaining a route for the emergence of early metalloporphyrins." *ChemBioChem* 23.8 (2022): e202200013.

3. Dagar, Shikha et al. "Nonenzymatic template-directed replication using 2′-3′ cyclic nucleotides under wet-dry cycles." *bioRxiv* (2022).

COPYRIGHT PERMISSIONS

Geochemical influences on nonenzymatic oligomerization of prebiotically relevant cyclic nucleotides

Shikha Dagar, Susovan Sarkar and Sudha Rajamani

RNA published online March 23, 2020

Nonenzymatic template-directed replication using 2′-3′ cyclic nucleotides under wet-dry cycles

Shikha Dagar, Susovan Sarkar, Sudha Rajamani*

Department of Biology, Indian Institute of Science Education and Research, Pune 411008, India

*E-mail: srajamani@iiserpune.ac.in
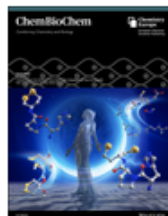
Thank you for your order!

Dear Shikha Dagar,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Shikha Dagar
Order Date:	Jul 14, 2022
Order Number:	5347740683521
Publication:	ChemBioChem
Title:	Porphyrin in Prebiotic Catalysis: Ascertaining a Route for the Emergence of Early Metalloporphyrins**
Type of Use:	Dissertation/Thesis
Order Ref:	20152012_CBC
Order Total:	0.00 USD

View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

VITA

Name: Shikha Dagar

Address: Department of Biology,

IISER-Pune, Dr. Homi Bhabha Road,

Ward No. 8, NCL Colony, Pashan,

Pune, Maharashtra – 411008.

Email address: shikha.dagar@students.iiserpune.ac.in

Education: B. Sc. (Hons.) in Chemistry, 2015

M.Sc., Biological Sciences, IISER-Pune, 2022

Ph.D., Biological Sciences, IISER-Pune, 2022