

# **Self-Assembly of nucleotides and its implications for prebiotic chemistry**

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

Shivam Tikoo



Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,  
Pashan, Pune 411008, INDIA.

Date: 10 April, 2023

Under the guidance of

Supervisor: Dr. Sudha Rajamani,

Affiliation of Supervisor

From May 2022 to Mar 2023

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

# Certificate

This is to certify that this dissertation entitled ‘Self-Assembly of nucleotides and its implication on prebiotic chemistry’ towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Shivam Tikoo at Indian Institute of Science Education and Research under the supervision of Dr. Sudha Rajamani, Department of biology, during the academic year 2022-2023.



Dr. Sudha Rajamani

Committee:


Dr. Sudha Rajamani

Dr. Gayathri Pananghat

This thesis is dedicated to my parents

# Declaration

I hereby declare that the matter embodied in the report 'Self-Assembly of nucleotides and its implication on prebiotic chemistry' are the results of the work carried out by me at the Department of Biology, Indian Institute of Science Education and Research, Pune, under the supervision of Dr. Sudha Rajamani and the same has not been submitted elsewhere for any other degree.

A handwritten signature in black ink that reads "Shivam Tikoo". The signature is written in a cursive style and is underlined with a single horizontal stroke.

Shivam Tikoo

Date: April 10, 2023

# Table of Contents

Declaration .....	iv
Acknowledgments .....	x
Contributions .....	x
Chapter 1 Introduction .....	1
Chapter 2: The Self-Assembly of nucleotides.....	11
Chapter 3: Implications of self-assembly on prebiotically relevant processes .....	40
Chapter 4: Conclusion and outlook .....	50

## List of Tables

Table 1: The shift in NMR signals for protons of various nucleotides (5' NMPs) at different concentrations in D<sub>2</sub>O.

Table 2: Summary of reactants used and their concentrations in a typical template-directed replication reaction.

## List of Figures

Figure 1: Comparison between the modern central dogma of life and the RNA World hypothesis.

Figure 2: Depicting different stages of a putative RNA World.

Figure 3: Proposed models (classical model, the ribose-centric model and the polymer fusion model) for the formation and emergence of the first genetic material on the prebiotic Earth.

Figure 4: The graph depicts the Rayleigh scattering for all the 5' NMPs.

Figure 5: Upper panel: a typical pyrene spectrum in methanol with I<sub>1</sub> and I<sub>3</sub> peaks labelled. The lower panel shows pyrene ground state monomer (P) interacting with an excited state molecule (P\*) to form an excimer (PP\*).

Figure 6: The plots show the normalized pyrene emission spectra at different concentrations of different 5' NMP solution.

Figure 7: The scatter plot shows pyrene I<sub>1</sub>/I<sub>3</sub> ratio (left panel) and pyrene I<sub>EX</sub>/I<sub>1</sub> ratio for different 5' NMPs at various concentrations.

Figure 8: The proposed model showing the partition comparison of purines and pyrimidines (left panel). The right panel shows how the presence of metal ions like Na<sup>+</sup> are known to induce and aid the formation of supramolecular structures of GMP.

Figure 9: Temperature-dependent change in the micropolarity of self-assembled structures of 5' NMPs. The scatter plot shows pyrene I<sub>1</sub>/I<sub>3</sub> ratio for purines (left panel) and pyrimidines (left panel).

Figure 10: Temperature-dependent change in micropolarity of self-assembled structures of 5' NMPs. The scatter plot shows pyrene I<sub>EX</sub>/I<sub>1</sub> ratio for purines (left panel) and pyrimidines (left panel).

Figure 11: Concentration-dependent change in anisotropy of self-assembled structures of 5' NMPs.

Figure 12: The plots show normalized pyrene emission spectra at different concentrations of AMP solution in the absence and in the presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride).

Figure 13: The plots show the normalized pyrene emission spectra at different concentrations of GMP solution in the absence and in presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride).

Figure 14: The plots show the normalized pyrene emission spectra at different concentrations of CMP solution in the absence and in presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride).

Figure 15: The plots show the normalized pyrene emission spectra at different concentrations of UMP solution in the absence and in presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride).

Figure 16: The scatter plot shows pyrene  $I_1/I_3$  ratios of indicated 5' NMPs in the presence of various metal ions.

Figure 17: The scatter plot shows pyrene  $I_{Ex}/I_1$  ratios of indicated 5' NMPs in the presence of various metal ions.

Figure 18: The plots show the normalized pyrene emission spectra at different concentrations of 5' NMP solution in the absence and in presence of 18% w/v PEG.

Figure 19: The scatter plot shows pyrene  $I_1/I_3$  ratios of indicated 5' NMPs in the presence of PEG.

Figure 20: The structures of various 5' NMPs (i.e., AMP, GMP, CMP and UMP) with labelled proton numbers).

Figure 21:  $H^1$  NMR characterization of AMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 22: The zoomed version of Figure 21 showing  $H^1$  NMR spectrum of AMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 23: The  $H^1$  NMR spectrum of GMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 24: The zoomed version of Figure 23 showing  $H^1$  NMR spectrum of GMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 25: The  $H^1$  NMR spectrum of CMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 26: The zoomed version of Figure 25 showing the  $H^1$  NMR spectrum of CMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 27: The  $H^1$  NMR spectrum of UMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 28: The zoomed version of Figure 27 showing the  $^1\text{H}$  NMR spectrum of UMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

Figure 29: This Figure shows 2D NOESY spectrum for CMP at 50mM concentration recorded in  $\text{D}_2\text{O}$ . H22-----H23, H22-----H7, H23-----H5 are the proton interactions that are clearly visible in this spectrum.

Figure 30: This Figure shows 2D NOESY spectrum for AMP at 50mM concentration recorded in  $\text{D}_2\text{O}$ . H10-----H7 H10-----H2 H7 -----H2 are the proton interactions clearly visible in this spectrum.

Figure 31: This Figure shows 2D NOESY spectrum for GMP at 50mM concentration recorded in  $\text{D}_2\text{O}$ . H7-----H10, H10-----H5, H5 ----- H7 are the proton interactions visible in this spectrum.

Figure 32: This Figure shows 2D NOESY spectrum for UMP at 50mM concentration recorded in  $\text{D}_2\text{O}$ . H16-----H5, H16-----H3, H16-----H7 are the proton interactions visible in this spectrum.

Figure 33: The illustration depicts how the presence of metal ions can induce the self-assembly of 5' NMP (this is showing GMP as an example in this context).

Figure 34: The illustration depicts how nucleotides could have driven the solubilization of hydrophobic molecules (peptide clusters are shown here as an example).

Figure 35: Schematics showing FDA absorption assay to check for the solubilizing capabilities of 5' NMPs.

Figure 36: The solubilization of FDA in the presence of different 5' NMPs at various concentrations.

Figure 37: Primer sequence and template sequence used in the nonenzymatic template replication. Template has highlighted base to which incoming nucleotide base pairs.

Figure 38: (A) Urea-PAGE representative gel image for control vs test reactions. At specific time points, the samples were taken from the reaction and were analyzed to eventually calculate the extension rate. Rate graph for (B) Control and (C) Test (BPY).

Figure 39: The zoomed in view of 5' NMP -BPY, BPY and 5' NMP absorbance spectra for various systems (A) AMP-BPY, BPY, AMP (B) CMP-BPY, BPY, CMP (C) GMP-BPY, BPY, GMP (D) UMP-BPY, BPY, UMP.



## **Abstract**

The self-assembly of nucleotides into RNA is believed to have been a crucial step in the emergence of cellular life. RNA, is a molecule which has a capability to store genetic information as well as catalyse chemical reactions that are essential for cellular processes. The ability of nucleotides to self-assemble would have played an essential role in the emergence of RNA on the prebiotic Earth. The presence of three subunits (a ribose sugar, a nucleobase, and a phosphate group) in its structure would have allowed nucleotides to interact in a variety of different ways. For example, the nucleobases can form hydrogen bonds with one another, creating base pairs that provide the structural basis for the double helix of DNA. Similarly, the phosphate groups can form bonds with the ribose sugars of adjacent nucleotides, creating the backbone of the RNA molecule. The self-assembly of nucleotides is a dynamic process that involves the interaction of multiple molecules and complex chemical interactions. Understanding this process is crucial in unravelling the shift from non-living matter to living organisms on Earth.

In this thesis, we have undertaken a comprehensive study of nucleotide self-assembly under a range of conditions and molecules. We investigated the self-assembly behaviour of different nucleoside monophosphates with respect to a change in the concentration, temperature and in the presence of prebiotically relevant co-solutes. Our results have demonstrated that the self-assembly of nucleotides is a highly responsive process that is influenced by environmental factors and strongly depends on the nucleobase present. We also studied the effect of this phenomenon on two prebiotically relevant processes: template-directed replication reactions and solubilization of hydrophobic molecules. Overall, our study has provided important insights into the complex and dynamic process of nucleotide self-assembly. By characterizing the factors that influence self-assembly, we hope to gain a deeper understanding of the origins of life and the fundamental processes that govern the behaviour of complex biological systems.

## **Acknowledgments**

I am deeply grateful to all the people who have contributed to the successful completion of this thesis. Without their support and guidance, this work would not have been possible.

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Sudha Rajamani. Their constant support, encouragement, and expertise were instrumental in shaping this project. I would like to thank my thesis advisor Dr. Gayathri Pananghat for her valuable inputs and support. Their valuable feedback and insights helped me to navigate through complex challenges and arrive at meaningful conclusions. I am also grateful to all the lab members who supported and encouraged me throughout this journey. Their input, discussions, and suggestions were critical in helping me to complete this work to the best of my abilities. The collaborative and supportive atmosphere of the lab was essential to my personal and professional growth. I would also like to acknowledge the support of the academic, technical, and non-teaching staff of the department. Their assistance and cooperation ensured that my research progressed smoothly and efficiently.

Finally, I would like to express my heartfelt gratitude to my family and friends. Their unwavering love, encouragement, and moral support gave me the strength and motivation to overcome obstacles and complete this thesis. Their presence in my life has been a constant source of joy and inspiration.

## Contributions

<b>Contributor name</b>	<b>Contributor role</b>
Shivam Tikoo	Conceptualization Ideas
Shivam Tikoo	Methodology
Shivam Tikoo	Software
Dr. Sudha Rajamani	Validation
Shivam Tikoo	Formal analysis
Shivam Tikoo	Investigation
Dr. Sudha Rajamani	Resources
Shivam Tikoo	Data Curation
Shivam Tikoo	Writing - original draft preparation
Dr. Sudha Rajamani	Writing - review and editing
Dr. Sudha Rajamani	Visualization
Dr. Sudha Rajamani	Supervision
Dr. Sudha Rajamani	Project administration
Dr. Sudha Rajamani	Funding acquisition

## **Chapter 1 Introduction**

Extant life is sustained by highly specialized biomolecules that maintain a far-from-equilibrium state. These complex biomolecules though present in extant biology were scarce on prebiotic Earth. The origins of life research area are a field of science that investigates the emergence of cellular systems from simple chemical and geological processes. This field covers a wide range of topics, including the formation of organic molecules, the evolution of self-replicating systems, and the emergence of cellular life. The central hypothesis in this field proposes that life evolved through chemical evolution, where increasingly complex molecules formed from simpler compounds, leading to the emergence of self-replicating systems and the first living organisms. This chapter will discuss various hypotheses and studies regarding how simple precursor molecules might have formed on early Earth, followed by the formation of complex biomolecule precursors and how the interaction between such prebiotic molecules might have aided the emergence of a primitive cell (protocell).

### **Origin of chemical precursors**

In the prebiotic era, given the implausibility of complex protein machinery, which aids the formation of molecules in biology, the formation of prebiotically relevant molecules is thought to have been nonenzymatic and would have been mainly aided by the surrounding environmental conditions. Towards this, there are studies debating the formation of prebiotically relevant molecules (such as nucleobases, amino acids, polyaromatic hydrocarbons (PAHs), porphyrin-like molecules, etc.) on the early Earth vs in extra-terrestrial regimes, where in the latter case, they were possibly delivered to the Earth via exogenous delivery by meteors, asteroids, etc (Ferris, Sanchez, and Orgel 1968; Oró and Kamat 1961; Schlesinger and Miller 1983). Irrespective of this, the reactions leading to the formation of such prebiotic molecules under early Earth conditions have also been demonstrated; the potentially more tangible alternative that allows us to understand how non-life might have transitioned to life on the early Earth. In the pioneering Urey-Miller experiments, the formation of biologically relevant molecules such as amino acids were shown starting

from simpler chemical molecules like water (H<sub>2</sub>O), methane (CH<sub>4</sub>), ammonia (NH<sub>3</sub>) and hydrogen (H<sub>2</sub>), under simulated early Earth conditions (Miller 1953). This experiment provided a proof-of-principle that early Earth conditions can be conducive towards the abiotic formation of biologically relevant molecules. Inspired by this, other studies have demonstrated the formation of molecules such as amphiphiles and sugars in Fischer-Tropsch like-synthesis and formose reaction, respectively (Oro 1995; Butlerow 1861). This further strengthened the abiotic origins of life hypothesis.

As indicated earlier, another source of these molecules could have been exogenous delivery. The cosmic inventory also has been shown to have an array of complex organic molecules. Meteorites have been found to contain purine and pyrimidine nucleobases, specifically guanine, adenine, cytosine, uracil, and thymine. (Oba et al. 2022). Another set of relevant molecules that are shown to be formed in extraterrestrial regimes and are detected in the meteoritic samples are polycyclic aromatic hydrocarbons (PAHs) (Groen et al. 2012). PAHs are abundant in the interstellar medium, extragalactic regions, protoplanetary disks, and other objects within our solar system, making them easily accessible in the primordial soup. Within the soluble organic phase of carbonaceous meteorites, various PAHs, including pyrene and fluoranthene, as well as oxidized aromatic species such as 9-fluorenone, 9-anthrone, 9,10-anthraquinone, and phenanthrene Dione, have been detected. Additionally, significant amounts of kerogen-type material, primarily composed of polymerized aromatics, have also been found (Ashbourn et al. 2007). It is known that PAHs can stack with an intermolecular separation distance of 0.34 nm, which is the same distance as nucleotides in RNA and DNA. This observation led to the 'PAH World Hypothesis', which proposes that PAHs that were abundant on early Earth and especially due to their stacking ability, would have preceded and aided the formation of the first genetic material, which is thought to be RNA (the RNA World hypothesis). Once these complex molecules are present on Earth, these molecules interact with each other under various conditions and give rise to emergent properties. This gives us a way to think about how the transition from simple to complex molecules could have come about prior to the emergence of early life.

## **RNA World Hypothesis**

According to the well-explored 'RNA world' hypothesis, RNA has conjectured the status of the first genetic polymer that emerged on prebiotic Earth during the evolution of life (Robertson and Joyce 2012). This is because it has the ability to serve as both a genetic polymer and a catalyst. So, RNA is considered to be a prime candidate for prebiotic information polymer. This hypothesis was further supported by the discovery of ribosome which contains catalytic RNA in its core. For the emergence of 'RNA World' abiotically, there is a requirement of nonenzymatic formation of monomeric substrates i.e., nucleotides, their polymerization to form RNA strands and replication of these RNA strands to propagate its information to the next generation, thus rendering it capable to undergo Darwinian evolution. Some progress has been made in all these stages of the putative 'RNA World', which will be discussed in the next part of this chapter.

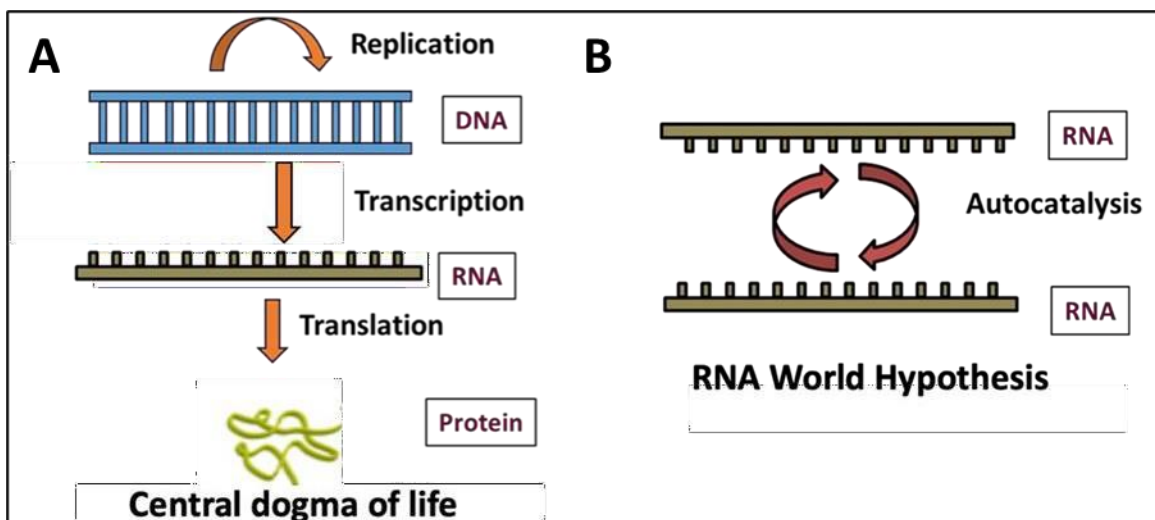


Figure 1: Comparison between the modern central dogma of life (A) and the RNA World hypothesis (B). As shown in the modern central dogma of life, DNA usually acts as a genetic material, protein acts as a catalyst and RNA acts as a bridge between the two, translating information from genetic information to proteins. In a putative RNA World as envisaged in the 'RNA World hypothesis', RNA can act as both genetic information as well as a catalyst Adapted from (Gilbert, W., Nature, 1986, Zaug, A., and Cech, T., Science, 1986)

### Nucleotides on early Earth

Nucleotides are the fundamental components of informational polymers, and determining their origins on the early Earth can provide insight into the first cellular systems. One proposed pathway involves the chemical reaction between molecules such as ammonia, methane, and water, in the presence of energy sources such as lightning, UV radiation, or volcanic activity (Odom et al. 1979; Schlesinger and Miller

1983; Yadav, Kumar, and Krishnamurthy 2020). In this scenario, the energy would drive the formation of more complex organic molecules, including nucleotides, through a series of chemical reactions (predominantly condensation reactions). Experiments conducted under laboratory simulated conditions have demonstrated that nucleotides can indeed form under early Earth conditions, lending support to this hypothesis. Another proposed pathway for the formation of nucleotides involves the delivery of prebiotic organic molecules to Earth via comets, meteorites, or other extra-terrestrial sources (Oba et al. 2022). These molecules could have also reacted with each other on Earth's surface or in the ocean, ultimately leading to the formation of nucleotides.

Extant biology utilizes nucleotides which are made up of three components mainly ribose sugar, phosphate group and nucleobase. Generally, nucleotides have three major components: a recognition unit [RU] which helps in hydrogen bonding with the other strand, a trifunctional connector [TC, like a ribose moiety], and an ionizable linker [IL, like the phosphate backbone] (Hud et al. 2013). Ribose and a phosphate group are utilized as TC and IL, respectively, to form ribose-phosphate backbone in contemporary biology. The major RUs in extant biology are the nucleobases adenine (A), guanine (G), cytosine (C), uracil (U) and thymine (T), which make up the genome of extant biology. In prebiotic chemistry, all these components were possibly formed abiotically, due to reactions based solely on the molecular properties of the constituting molecules, thereby resulting in a plethora of similar molecules. This has been shown in the form of threose ribonucleic acid (TNA), where ribose is replaced by threose (Orgel 2000). Researchers have also changed the ribose-phosphate backbone to N-(2-aminoethyl)-glycine (AEG) connected with peptide bond to form Peptide nucleic acid (PNA) (Nelson, Levy, and Miller 2000). Such PNAs, due to the absence of the negatively charged phosphate group, show greater binding with the ssDNA or RNA, to form PNA/DNA or PNA/RNA as compared to the DNA/DNA or RNA/RNA duplexes. As for the nucleobase part, theoretically, there are more than 80 prebiotically plausible nucleobases, among which few alternate nucleobases are explored as prebiotically relevant RUs due to their capability to hydrogen bond in a similar way as that of extant nucleobases. Previous studies have utilized 2,4,6-triaminopyrimidine (TAP), cyanuric acid (CA) and melamine (MA) (a triazine) as alternate nucleobases to form genetic polymers (Hud et al. 2013).

Next, a trifunctional connector (TL) is the unit which connects RU and IL. These in extant biology are deoxyribose or ribose sugar. However, ribose sugars are complex and less stable molecules thus it is logical to assume that other plausible sugar moieties could have been present before transitioning to ribose. For e.g., these can be replaced by molecules like threose etc. Further, for the Ionizable linkers (IL), extant biology utilizes phosphate as the backbone of the RNA/DNA. These are replaceable and studies have shown the use of phosphate derivatives to form the backbone. However, the major challenge for phosphate groups is their instability under early Earth conditions, which are thought to have been harsh (high temperature). All of this necessitates the requirement of a niche that is conducive for such condensation reactions, but at the same time need to be less detrimental towards the resultant polymers. Therefore, there can be many possible nucleotide or nucleotide-like molecules which can form the genetic information of a protocell. Once these nucleotides form, they need to form a short stretch of genetic polymer and replicate in order to propagate its information to the next generation.

### **Nucleotides in information polymer (Nonenzymatic polymerization and replication)**

In extant biology, the primary function of nucleotides is to carry hereditary information and facilitate Darwinian evolution, which is orchestrated by several highly precise protein molecules. The information polymer on the prebiotic Earth would have however had to catalyze its own replication. This is because, on the prebiotic Earth, the formation of enzymes would have been non-trivial (the chicken and egg paradox), so for RNA to replicate, it would have had to be enzyme free. The first step in all of this is forming an information polymer by nonenzymatic oligomerization, Various studies have shown the formation of small stretches of oligonucleotides with canonical as well as non-canonical nucleotides, using nonenzymatic oligomerization (Mungi et al. 2019; Dagar, Sarkar, and Rajamani 2020; Kaddour et al. 2018).

Researchers have tried to form stretches of RNA from various types of monomeric nucleotides. Studies have shown the formation of stretches of RNA by mimicking prebiotic conditions. There have also been reports where they have used different metal ions which can catalyze the polymerization processes (Ferris 1993) of which  $Mg^{2+}$  is one of more efficient in polymerization and of oligomerisation RNA. The use of activated nucleotides has also resulted in the formation of oligomers (Dagar,



Sarkar, and Rajamani 2020; Burcar et al. 2015). However, these activated nucleotides are unstable at high temperatures and high salt concentrations.

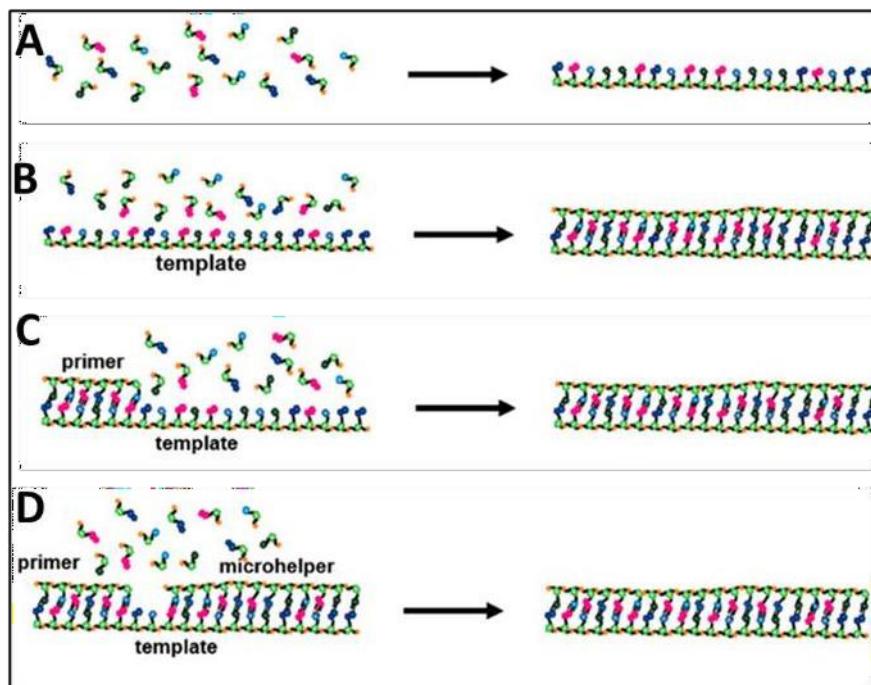


Figure 2: Depicting different stages of a putative RNA World. The first one is nonenzymatic oligomerization of monomeric nucleotides to form RNA strands. The second is template-directed oligomerization where addition of nucleotides occurs in a sequential manner to copy the genetic information of a template RNA strand. The third is template-directed replication where nucleotides are sequentially added to a primer annealed to a template. The fourth is template-directed replication where nucleotides are added in between a primer and a downstream helper oligomer which are annealed to a template. Adapted from (Kaddour and Sahai 2014)

Unlike in extant biology once these oligonucleotides are formed, the information polymer needs to catalyze itself without any replication machinery. Non-enzyme-mediated RNA replication could serve as an intermediary step between the production of nucleotides in prebiotic conditions and the established RNA world, where RNA enzymes or ribozymes may have facilitated replication of early cells that had RNA genomes. In the field, researchers typically use 5' activated nucleotides (imidazole activated) which are incorporated against the respective templating base, to study the replication reaction. The activated nucleotides have a good leaving group, which makes it attack the nucleophilic NH<sub>2</sub> center on the 3'-NH<sub>2</sub> terminated primer, which leads to extension. This reaction gives us a plausible mechanism on how RNA could have replicated under prebiotic conditions (Szostak 2012). Substantial work has been done on this where the rate for replication has been

calculated depending on the leaving group(Izgu et al. 2015). Interestingly, most of these reactions are performed under a scenario of a controlled environment i.e., devoid of any background molecules that, could which might affect the reaction. This kind of situation fails to depict the complex chemical conditions that could have existed on the Earth before the emergence of life.

The primordial soup is thought to have consisted of an array of various chemical entities that could have acted as co-solutes. These co-solutes could have had a significant effect on prebiotic oligomerization and replication of early RNA. In this regard, one of the relevant studies from the lab investigated the effect of the presence of co-solutes such as: either Polyethylene glycol (PEG), 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) or both on RNA replication. The findings indicated that co-solutes led to a notable decrease in the incorporation of purine monomers opposite their corresponding template base. (Bapat and Rajamani 2015).

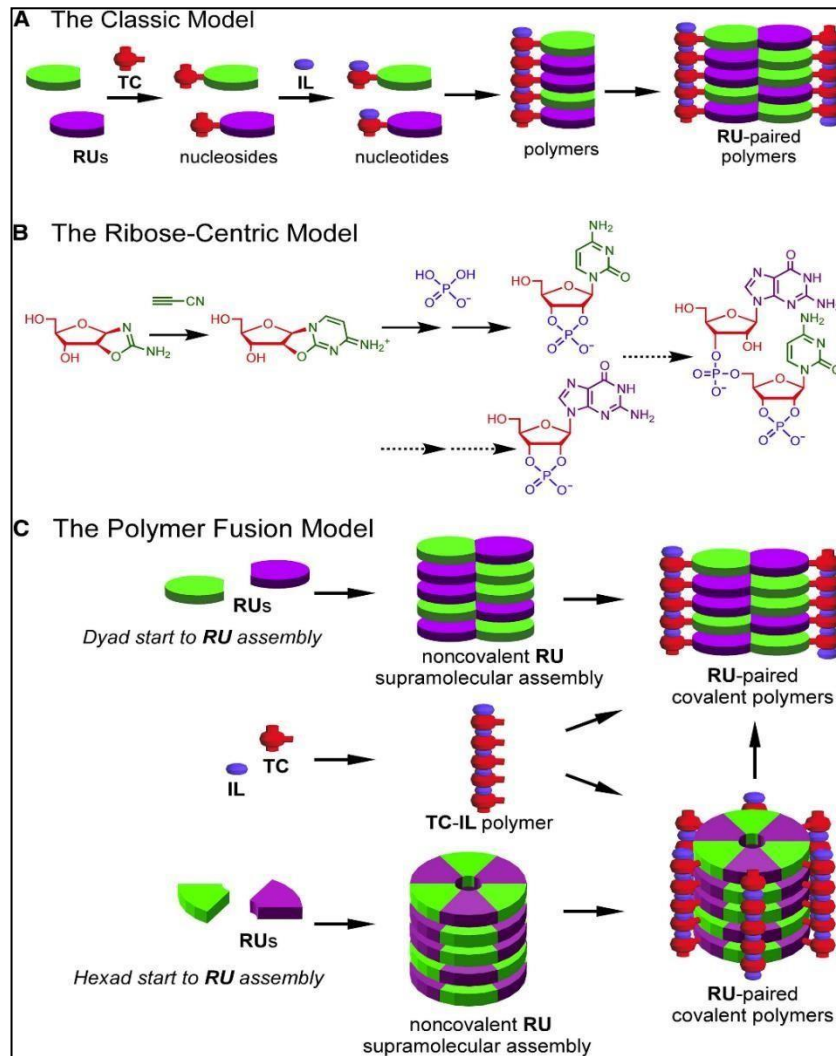


Figure 3: Proposed models for the formation and emergence of the first genetic material on the prebiotic Earth. A) according to the classical model, nucleotide is divided into three parts: the recognition unit (RU), trifunctional connector (TC), and ionic linker (IL) where all of them combine to form a nucleotide or protonucleotides (primitive form of nucleotides). These nucleotides/protonucleotides then condense abiotically to form nucleic acid. (B) The second model is ribose-centric model. As the name suggests, it proposes that the nucleobase is formed on a pre-existing sugar. These nucleosides then phosphorylate to form nucleotides which then condenses sequentially to form RNA strands. (C) The third model is the polymer fusion model. This model proposes that Recognition units (RUs I.e., nucleobases) then bind with each other to form hydrogen bonded dyads/tetrads/hexads which self-assemble to form long stacks. The polymeric backbone containing repeating units formed by simultaneous condensation of Trifunctional connectors (TCs) and ionized linkers (ILs) can then covalently fuse with these supramolecular assemblies to form polymers. Adapted from: (Hud et al. 2013)

Possible hypotheses for the formation and replication of nucleotides include the classic model, that involves sequential assembly of the ionic linker, recognition unit and trifunctional connector to produce nucleotides before polymerization to form RNA, the ribose-centric model, that which involves construction of the RU from a

pre-existing sugar and the polymer fusion model, which involves the fusion of RUs and TCs through covalent linking to form informational polymers (Hud et al. 2013).

## **Protocell**

Decoding the transition from complex chemical molecules to interacting entities in a cellular machinery is challenging. Protocells, which are minimalistic forms of cellular systems, possess three characteristic features: information polymer, compartmentalization, and the use of a minimal metabolic network. In principle, these can replicate, transfer genetic material, and sustain non-equilibrium conditions for cellular activities. Protocells are believed to have emerged through the self-assembly of simple molecules like lipids into vesicles, which can encapsulate other molecules such as RNA or peptides. This area of research is active, and scientists are exploring different ways in which protocells could have formed and evolved on early Earth, as well as their potential role in the origin of life on other planets and also for the design of synthetic life (Morowitz, Heinz, and Deamer 1988).

Protocell compartments, made from lipids or other molecules, are important for differentiation and protection of encapsulated material from dilution and other pathogenic molecules. Lipids, which are amphiphilic, can self-assemble into bilayer structures such as vesicles, creating simple protocells that can exchange materials with their environment. Single chain amphiphiles (SCA) are considered prime contenders for prebiotic compartments due to their simpler chemical structure, variety, and prebiotic availability to self-assemble into vesicles (Deamer 2016). SCA vesicles need to be dynamic and fluid to allow desired molecules to diffuse inside, unlike stable phospholipid vesicles.

The metabolism aspect is crucial for a protocell to convert nutrients into energy and useful molecules. Metabolic pathways must operate in a self-contained system without external enzymes. Simple pathways and molecules that can perform reactions with external energy need to be identified, as the metabolic pathways in extant biology are too complex for a prebiotic setting. Some molecules, such as PAHs and their derivatives, can capture light energy to generate higher chemical potential or ionic gradients, making them capable of mediating the conversion of lipid precursors into lipids (Mahajan et al. 2003).

The self-assembly of nucleotides is a complex and dynamic process that has significant implications for prebiotic chemistry. Nucleotides are the building blocks of nucleic acids, which are essential for the storage and transmission of genetic information in all living organisms. Understanding how nucleotides self-assemble is therefore crucial for unraveling the origins of life and the evolution of biological systems. Nucleotides can self-assemble through a variety of mechanisms, including electrostatic interactions, hydrogen bonding and hydrophobic interactions. These interactions leads to the formation of complex structures such as oligonucleotides, nucleic acid polymers, and even protocells. Overall, this thesis aims to explore the complex and dynamic nature of nucleotide self-assembly and its implications for prebiotic chemistry.

## Chapter 2: The Self-Assembly of nucleotides

Self-assembly is the phenomenon where the physical and chemical characteristics of specific types of molecules cause them to interact and form complex assemblies. These structures can vary from micron scale to nano scalar assemblies. Self-assembly process is thermodynamically driven, the system tries to increase its entropy and thus assemble into different structures. This process usually gives rise to emergent properties. In extant biology self-assembly is a very prominent phenomena and can be seen in cellular systems. The assembly of DNA duplex, secondary structure of RNA and proteins, are all dependent on self-assembly of molecules. Apart from cellular processes, researchers have used self-assembly in forming nanoparticles, metal organic framework (MOF) and nano tubes for facilitating various purposes like chemical sensing, drug delivery, etc (L. Zhou et al. 2014; Wong et al. 2005; Liang et al. 2017).

In prebiotic chemistry, self-assembly would have played a very important role in setting the early stages for the eventual emergence of cellular systems (Deamer et al. 2006). The most prominent self-assembled structure on early Earth would have been vesicles, formed from various Single chain amphiphiles (SCAs). Lipid molecules self-assemble and form a vesicular system which is an important part of the protocell as mentioned earlier. There have been extensive studies related to self-assembled vesicular systems to understand their dynamics and physicochemical properties. Another self-assembly process which is often overshadowed by lipids is the self-assembly of nucleotides. Nucleotides can self-assemble at high millimolar concentration, and this assembly is mostly facilitated by pi-pi interactions (A. Sigel, Operschall, and Sigel 2014; Stokkeland and Stilbs 1985). The monophosphate has better stacking capabilities than triphosphate because the extra negatively charged di-phosphate in the latter leads to repulsion between two nucleotides triphosphates. Prebiotically, monophosphates are thought to be more readily available as compared to triphosphates, whose aggregates can be affected in the presence of metal ions, high temperature etc. (H. Sigel 2004).

The most prominent nucleotide self-assembled structure that has been studied are GMP quartets/G-quartets. GMP's ability to form Hoogsteen base pairing leads to

formation of quartets. These quartets then can assemble on each other to form cylindrical structures. Previously, the formation of hydrogels was shown when amphiphilic molecules trap water molecules, making crosslinked networks. GMP has also been shown to readily form hydrogels and properties like temperature responsiveness, pH responsiveness can be modulated by addition of other nucleotides or changing the pH of the system or by addition of salts (Cassidy et al. 2014). Earlier studies have also utilized GMP/ Guanosine hydrogels to deliver drugs. nucleotides can also assemble into MOFs forming nanozymes (Liang et al. 2017). Self- assembly properties of other nucleotides have also been studied, but not in the context of origins of life. It is well known that self-assembled structures are usually dynamic and are in a state of equilibrium. The addition of different co-solute molecules or changes in the environmental conditions can, therefore, impact these assembled structures. Pertinently, the tendency to form self-assembled structures can also influence various processes by regulating the availability of constituting molecules. Thus, it is hypothesized that the presence of self-assembled structures (for example: self-assembled structures of nucleotides) would have influenced various prebiotically relevant processes. To further understand this and how the presence of various co-solutes (metal ions, PAH, PEG etc.) would have influenced this phenomenon, we developed assays to characterize and study the self-assembled structures of 5' NMPs. We also investigated how the presence of putative self-assemblies could impact prebiotically relevant processes such as enzyme-free template-directed replication reactions and solubilizing properties of 5' NMPs (Chapter 3).

## **Experimental section**

### **Structural analysis of self-assembled nucleotide structures to understand their dynamicity**

Self-assembled structures are usually dynamic. Even though it has been established that nucleotides can interact with each other in high millimolar regimes, the dynamicity of the system is still not fully known. In order to understand physicochemical properties, solvatochromic fluorescent probes such as pyrene can be utilized. Pyrene has a distinctive emission spectrum that displays five distinct peaks, and the proportion of peak 1 to peak 3 indicates the micro polarity of the

surroundings (Xu et al. 2018). The highest ratio is observed when the environment is polar, and it decreases as the environment's polarity decreases (such as when pyrene is incorporated into a hydrophobic interface). Our research aims to investigate the physicochemical properties of self-assembled structures of various 5' NMPs, in order to gain a better understanding of their impact on prebiotically relevant processes. To further confirm the presence of self-assembled structures, Rayleigh scattering experiments were also used. In order to comment upon the nature of assembly, anisotropy experiments were performed using pyrene that indicates its movement restriction in the presence of nucleotides or self-assembled structures. Further, to understand structurally how the resultant aggregates are interacting,  $^1\text{H}$  NMR and 2-D NOESY NMR were recorded, which can shed light on how two 5' NMP molecules are interacting. Once the physicochemical properties of such self-assembled structures were characterized, a change in the above-mentioned properties in the presence of various co-solutes such as metal ions and PEG was also investigated.

## **Materials and Methods**

The disodium salts of all four 5'-NMPs, viz. 5'-AMP, 5'-GMP, 5'-UMP, and 5'-CMP (all purity  $\geq 98\%$ ), pyrene, and analytical grade PEG 8000 were purchased from Sigma-Aldrich (Bangalore, India) and used without any further purification.

### **Sample preparation**

Nanopure water was used to prepare nucleotide stocks, whose concentrations were estimated using UV/VIS/Vis spectroscopy (UV/VIS-1800 UV/VIS/Vis spectrophotometer, Shimadzu Corp., Japan). Four different concentrations of the nucleotides viz. 0 mM, 10 mM, 50mM, 100mM, were used to record all the measurements.

### **Rayleigh scattering for nucleoside 5'-monophosphates (5' NMPs)**

The 5' NMP samples were excited with a light of wavelength 400 nm and the emission intensity was recorded at the same wavelength i.e., 400 nm. The emissions and excitation slits were fixed at 2 nm. The fluorescence intensity was measured for



0 mM of each 5' NMP solution till 300 seconds, after which appropriate volumes of the corresponding 5' NMP stock solution was added and mixed so that the final concentration was 10 mM. The reading was monitored for another 300 seconds so that the signal was stabilized. Following this, the same procedure was repeated to get the stabilized fluorescence intensity for 50 mM and 100 mM of final concentrations of the various 5' NMPs. All these reactions were carried out at room temperature.

### **Steady-state fluorescence analysis using pyrene**

The self-assembly behavior of different NMPs was studied using a solvatochromic fluorophore i.e., pyrene. To carry out the experiment, 2  $\mu\text{l}$  of a 200  $\mu\text{M}$  solution of the dye was mixed with 100  $\mu\text{l}$  of the sample, resulting in a final concentration of 5  $\mu\text{M}$ . The mixture was then incubated at 25°C while being shaken at 400 rpm for 20 minutes to allow the dye to incorporate into the hydrophobic regions of the self-assembled structures. Pyrene was excited at 335nm and the emission spectra was taken from 350 to 600nm. The fluorescence readings were taken at a 90° angle using a FluoroMax 4 fluorescence spectrophotometer (made by Horiba Jobin Vyon) with a 150 W CW ozone-free xenon arc lamp. The excitation and emission slit width were maintained at 2 nm, while for anisotropy measurements, they were kept at 3 nm. Lamp intensity variations were monitored and corrected as needed.

### **Temperature-dependent fluorescence analysis using pyrene**

For temperature-dependent fluorescence studies, either 10 mM or 50 mM of corresponding 5' NMP solution was preheated for 15-20 min at the required temperature. The UV/Vis cuvette (obtained from Helma) and the UV/Vis cuvette holder were separately incubated at the same temperature. After the incubation, the solution was transferred to the UV/Vis cuvette and the fluorescence emission spectrum was recorded for pyrene. All other parameters such as pyrene concentration (5  $\mu\text{M}$ ) and slit width (2 nm) were kept constant.

### **Self-assembly in presence of different co-solutes**

The physicochemical properties of self-assembled structures of 5' NMPs were studied using pyrene in the presence of various metal ions (KCl, NaCl and  $\text{MgCl}_2$ ). In a typical reaction, 100 mM for KCl/NaCl or 50 mM for  $\text{MgCl}_2$  was added to the 5'

NMP solution with varying concentrations (to a final concentration of 0.1 mM, 1 mM, 10 mM, 50 mM and 100 mM). As mentioned above, pyrene at a final concentration of 5  $\mu$ M was added to each experiment and fluorescence readings were recorded at room temperature. All other experimental parameters were kept same as that of steady state fluorescent measurements using pyrene.

### **NMR data acquisition**

NMR experiments data acquisition was performed on Bruker 400 MHz NMR spectrometer, furnished with BBI Broadband Inverse Double Channel Probe.  $^1\text{H}$  NMR experiments were recorded for different concentrations of nucleotides viz. 10 mM, 50 mM and 100 mM, dissolved in  $\text{D}_2\text{O}$ . 2D NMR experiments were recorded at 50 mM for all the nucleotides.

### **Results and Discussion**

#### **Rayleigh scattering for Nucleoside 5'-monophosphates (5' NMPs)**

In order to get an initial confirmation that nucleotides do form supramolecular structures, Rayleigh scattering was used. 5' NMP samples were excited at 400 nm, and emission fluorescence intensity was monitored at the same wavelength i.e., 400 nm. It was made sure that nucleotides do not absorb at that wavelength. As can be seen in Figure 4, an increase in scattering intensity was observed with an increase in the concentration of the sample over time. The concentrations of 5' NMPs used were 0 mM (just nanopure water), 10 mM, 50mM and 100mM. After addition, each concentration was given some time to stabilize after which we didn't see any fluctuations in the intensity for each concentration. This experiment indicated that there is a formation of higher order structures, which could be probably due to the self-assembly of nucleotides given that they have a heterogenous base that can

facilitate this.

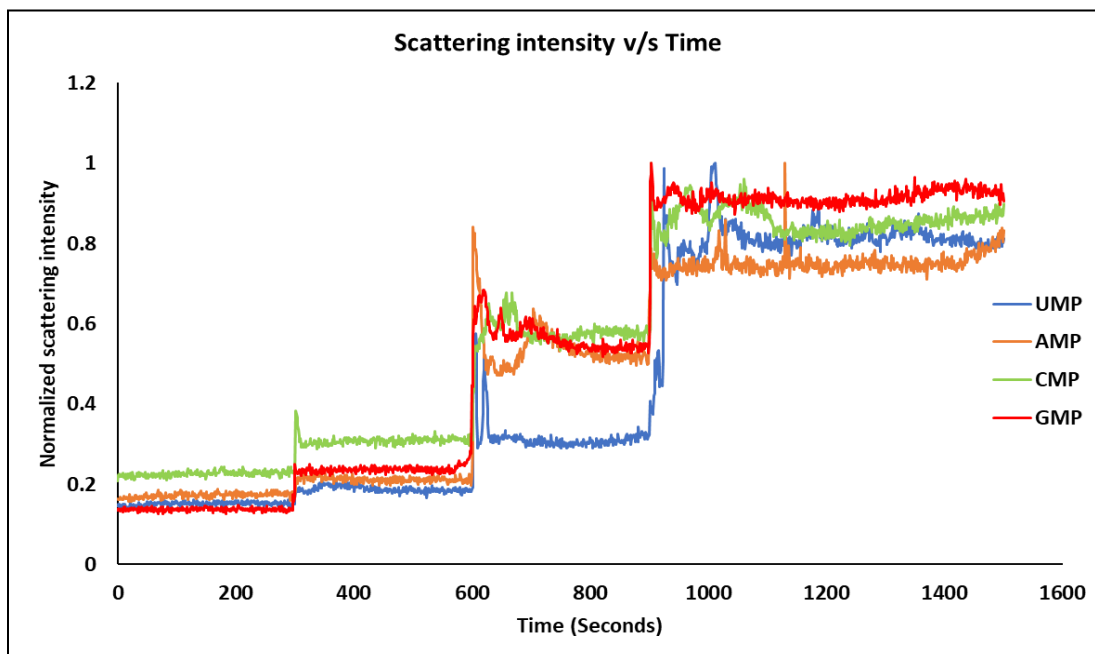


Figure 4: The graph depicts the Rayleigh scattering for all the 5' NMPs. Nucleoside monophosphate samples were excited at 400 nm, and emission fluorescence intensity was monitored at the same wavelength. The X-axis represents time (in seconds) and the Y-axis indicates normalized scattering intensity. 5' NMPs were added such that the final concentration was 10mM, 50mM and 100 mM after 300, 600 and 900 seconds, respectively. There was a clear increase in the scattering intensity with the increase in the 5' NMP concentration, which can indicate the formation of higher-ordered structures.

### Pyrene assay for 5' nucleoside monophosphate

As previously stated, pyrene is a polycyclic aromatic hydrocarbon (PAH) that has a unique emission spectrum. It exhibits five distinct peaks, and the ratio of the first peak to the third peak ( $I_1/I_3$ ) is used to determine the micro polarity of its surroundings. This ratio is highest in polar environments and decreases as the environment becomes less polar (such as when pyrene is partitioned into a hydrophobic interface). The highest  $I_1/I_3$  ratio for pyrene was observed in water. Pyrene also demonstrates excimerization at approximately 460 nm, which is measured by the  $I_{ex}/I_1$  ratio (the emission fluorescence intensity at 460 nm divided by the intensity of the first peak). The formation of pyrene excimers has been the focus of numerous studies, which have proposed various models for the process. Some models suggest that excimerization occurs when the excited state pyrene ( $P^*$ )

monomer collides with the ground state pyrene (P) monomer, implying that diffusion plays a significant role in the process. Other studies suggest that the static contact between the P and P\* monomers, when they are in close proximity, leads to excimer formation through pyrene aggregation in the hydrophobic core (as shown in Figure 5). However, in all of these scenarios, contact between the P and P\* monomers results in excimerization. (Barenholz et al. 1996).

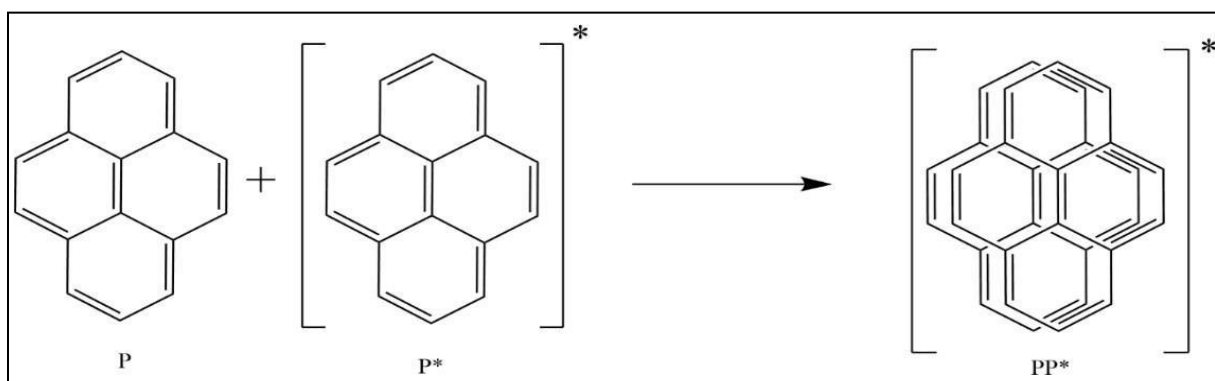
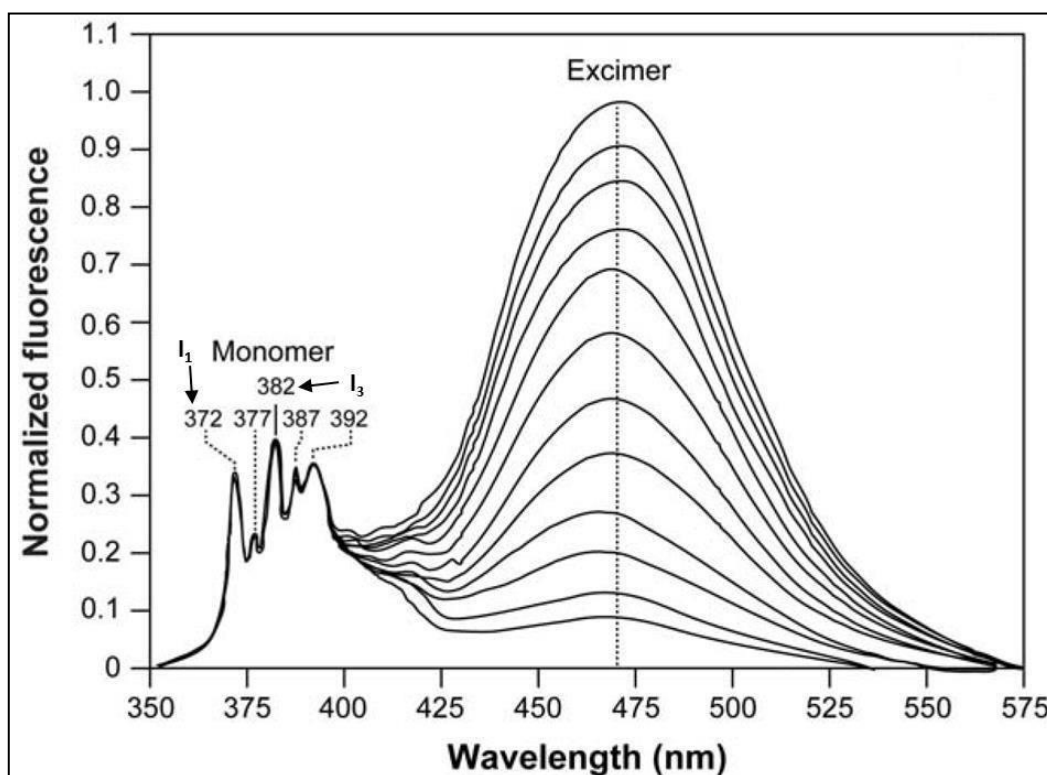


Figure 5: Upper panel shows pyrene spectrum in methanol with  $I_1$  and  $I_3$  peaks labelled. Increasing concentration leads to the formation of excimerization at around 460-470nm. Adapted from Zaragoza-Galán, et al. *Molecules* 19.1 (2013): 352-366. The lower panel shows pyrene ground state monomer (P) interacting with an excited state molecule (P\*) to form an excimer (PP\*).

In order to examine the characteristics of the self-organized structures pyrene emission spectra were obtained at different concentrations for each of the 5' NMPs and  $I_1/I_3$  and  $I_{Ex}/I_1$  ratios were quantified (Figure 6 and 7). As is evident in Figure 6, the pyrene fluorescence spectra are different for all 5' NMPs. As seen in Figure 7, a decrease in the  $I_1/I_3$ , indicating an increase in hydrophobicity with an increase in concentration, was observed for all the systems except GMP, which showed a peculiar result. Surprisingly, GMP first showed a decrease in the  $I_1/I_3$  due to an increase in hydrophobicity when the concentration was increased from 0 mM to 10 mM and 50 mM. However, after this there was a sudden increase in  $I_1/I_3$  at 100 mM. Similarly, for excimerization, an increase in  $I_{Ex}/I_1$  with increasing concentration was observed for all the 5' NMPs excepting for AMP and GMP. AMP did not show any significant change in  $I_{Ex}/I_1$ . In the case of GMP, a sudden decrease was observed at 100 mM. This can potentially be explained because of the partition of pyrene in a heterogenous population of different higher order structures, as GMP is known to form structures like tetrads and cylindrical tubes at high concentrations, both of which have been reported in previous studies. Therefore, there is a possibility that there may be two different populations of GMP aggregates present at higher concentrations, and one is favored over the other depending on the concentrations. At higher concentration of 100mM, the aggregate morphology is possibly making it inaccessible for pyrene to stack between them. Being pyrimidines, UMP and CMP have stacking efficiency that is not as prominent as the purines, leading to unstable and transient aggregates, which potentially is leading to higher excimerization values

as seen in Figures 6 and 7.

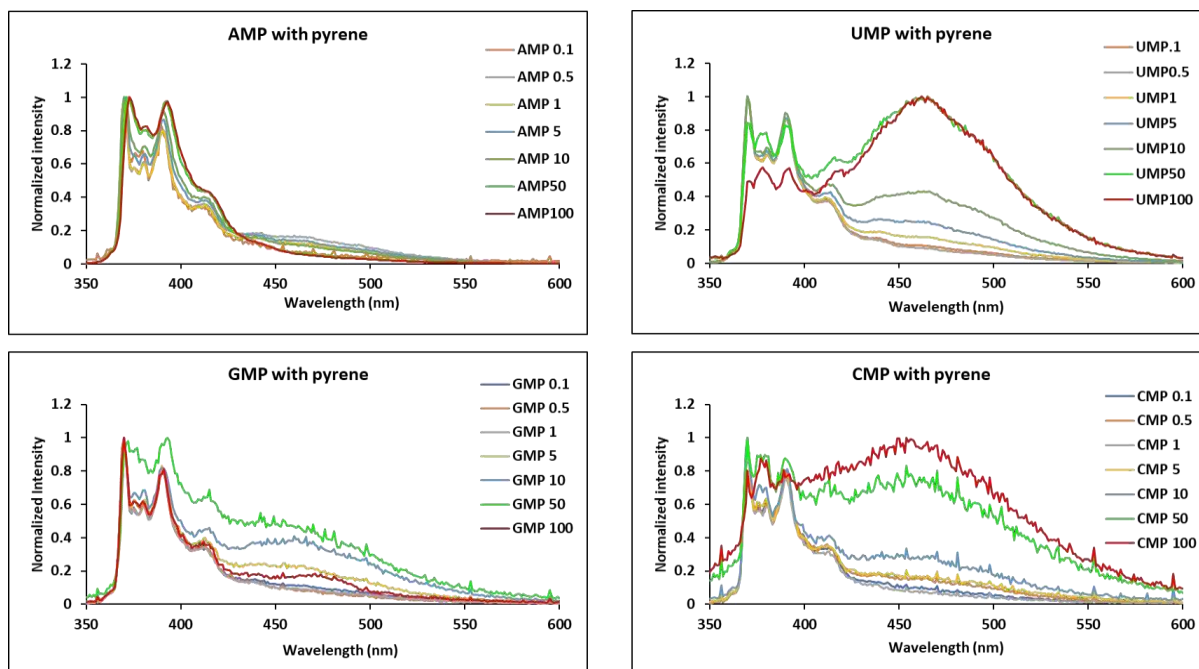


Figure 6: The plots show the normalized pyrene emission spectra at different concentrations of indicated 5' NMP solution. The Y-axis indicates normalized fluorescence intensity and X-axis indicates wavelength (in nm).

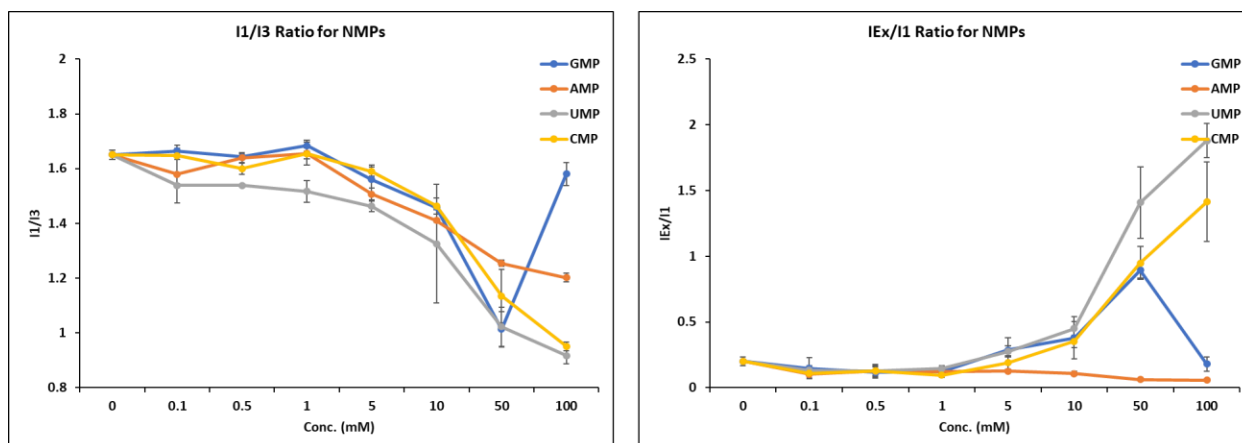


Figure 7: The scatter plot shows pyrene  $I_1/I_3$  ratio (left panel) and pyrene  $I_{Ex}/I_1$  ratio for different 5' NMPs at various concentrations. The X-axis indicates the concentration in mM and the Y-axis indicates the indicated ratios for the corresponding 5' NMPs. Different colors depict different 5' NMPs studied. N=3, error bar=SD.

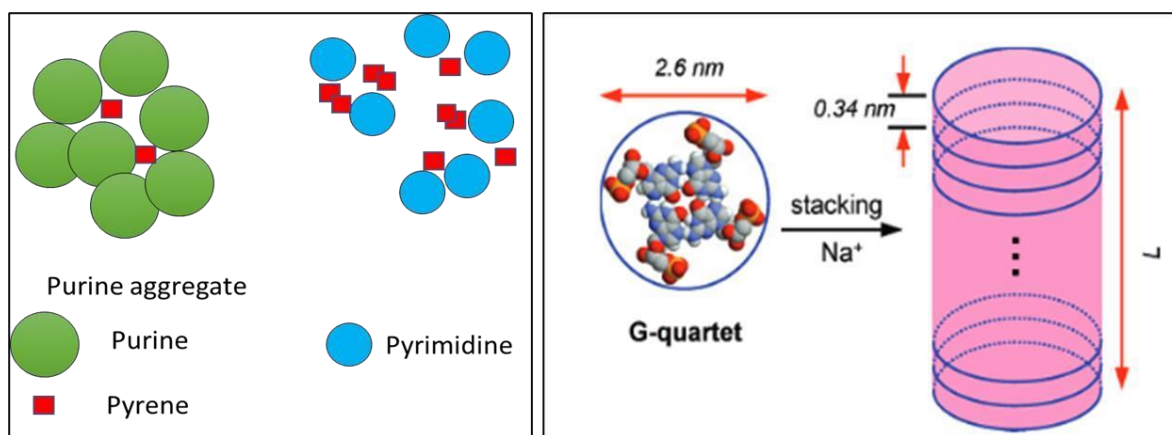


Figure 8: The proposed model showing the partition comparison of purines and pyrimidines. The presence of rigid self-assembled structures in the case of purines when compared to the dynamic and transient assemblies resulting in the case of pyrimidines, can potentially explain a lower excimerization in purines when compared to pyrimidines. The right panel shows how the presence of metal ions like  $\text{Na}^+$  are known to induce and aid the formation of supramolecular structures of GMP. Adapted from (Wong et al. 2005)

The major trend of decrease in hydrophobicity seen confirms the formation of aggregates as also indicated by an increase in Rayleigh scattering, while the excimerization depicts the dynamicity of these self-assembled structures. Thus, these results help in understanding the self-assembly and the dynamicity of the self-assembled structures of 5' NMPs at various concentrations. We propose that the purines are closely bound in their self-assembled structures, thereby not allowing the pyrene molecules to diffuse and interact with each other, leading to low or no excimerization unlike in pyrimidines (Figure 8). This also corroborates with the fact that purines have better capability to stack than pyrimidines do.

### Temperature dependent pyrene assay for nucleoside monophosphate

Next, we wanted to study the dynamics and the stability of the self-assembled structures of various 5' NMPs. For that, we investigated the temperature-dependent changes in the physicochemical properties of these structures. In a typical experiment, the  $I_1/I_3$  and  $I_{Ex}/I_1$  ratios of pyrene fluorescence was monitored for 50 mM of the corresponding 5' NMP solutions and at different temperatures (30°C, 40°C, 50°C and 60°C).

As observed in Figure 10, the  $I_1/I_3$  ratio for AMP remains unchanged over the investigated temperature range. For GMP, the  $I_1/I_3$  ratio stays the same at 30°C and 40°C, after which it decreases significantly to 50°C and 60°C. In the case of pyrimidines, i.e., CMP and UMP, the  $I_1/I_3$  ratio decreases with an increase in

temperature. As observed in Figure 10, there was an increase in the  $I_{Ex}/I_1$  i.e., excimerization in the case of AMP and GMP. This is expected as aggregate stability decreases with increase in temperature, making them more transient. This would result in more ready interactions between the pyrene molecules that could result in higher order structures, making a larger number of smaller hydrophobic aggregates. In the case of GMP, a drastic increase was observed in  $I_{Ex}/I_1$  at 60°C. This could be due to the formation of a different population of aggregates, wherein the smaller sized population is possibly getting favoured at higher temperatures. Contrarily, for CMP and UMP, excimerization decreased with an increase in temperature, which can be explained by less stacking capability of pyrimidines. These results helped us to understand the stability of aggregates over different temperatures. Overall, the stability of the aggregates decreased over higher temperature. AMP and especially GMP exhibited relative increase in excimerization at high temperatures, whereas in the case of UMP and CMP, the excimerization decreased relatively with an increase in temperature.

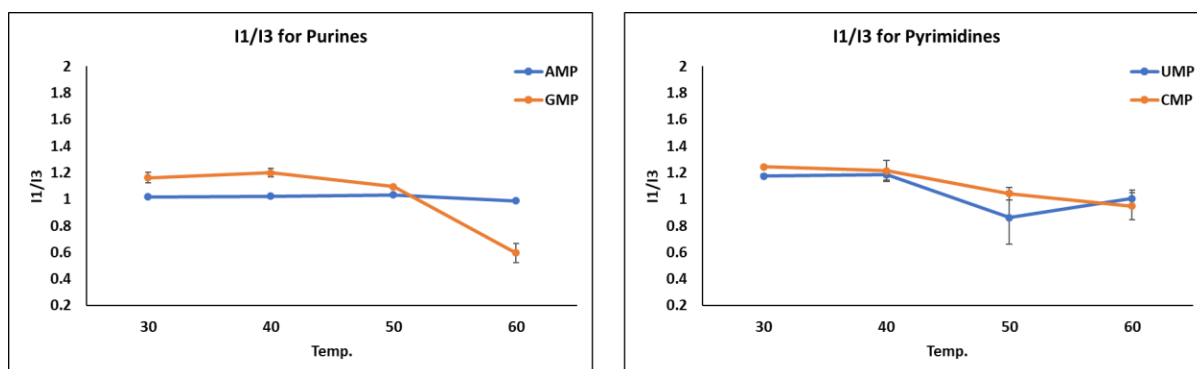


Figure 9: Temperature-dependent change in micropolarity of self-assembled structures of 5' NMPs. The scatter plot shows pyrene  $I_1/I_3$  ratio for purines (left panel) and pyrimidines (left panel). The X-axis shows the temperature in Celsius, and the Y-axis shows the indicated ratios for corresponding 5' NMPs. Different colors depict different 5' NMPs used. N=2, error bar=SD.



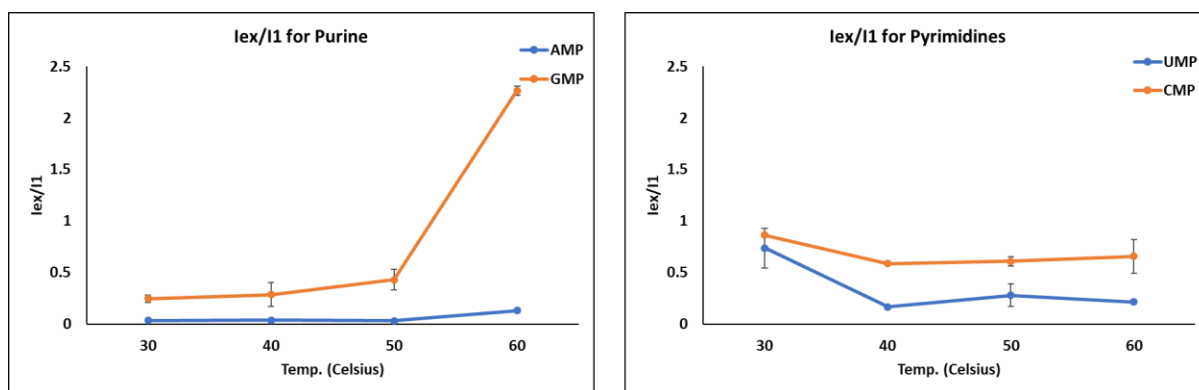


Figure 10: Temperature-dependent change in micropolarity of self-assembled structures of 5' NMPs. The scatter plot shows pyrene  $I_{EX}/I_1$  ratio for purines (left panel) and pyrimidines (left panel). The X-axis shows the temperature in Celsius, and the Y-axis shows the indicated ratios for corresponding 5' NMP. Different colors depict different 5' NMPs used. N=2, error bar=SD.

### Pyrene anisotropy for nucleoside monophosphate

Fluorescence anisotropy is the term used to describe the polarization of light that is emitted by a fluorophore. The degree of polarization of the emitted light is measured as anisotropy, which provides information about the rotational motion of the fluorophore and its local environment. Fluorescence anisotropy is used to study molecular-level motions. Fluorescent molecules are activated by plane-polarized light, and if the molecules are stationary, they continue to emit light in the same polarization plane. However, the light released will be in a different plane from that of the original excitation if the excited molecule spins or falls out of the plane of the polarized light while in the excited state. The intensity of the emission light can be measured in both the original vertical plane and the horizontal plane when utilizing vertically polarized light to stimulate the fluorophore. The degree to which the emission intensity shifts from a vertical to a horizontal plane is correlated with the fluorophore's mobility. pyrene fluorescence anisotropy was thus used to understand the pyrene movement in the resultant 5' NMP aggregates.

All the 5' NMPs were used in five different concentrations i.e., 0.1, 1, 10, 50, and 100 mM. We see a clear trend in the anisotropy of the 5' NMPs; As seen in Figure 11, UMP showed the highest anisotropy followed by CMP, GMP, AMP. This could be due to the entrapment of pyrene molecules in several smaller aggregates of UMP, thus restricting its movement, leading to high anisotropy. Based on these results, we

can also comment about the size of the resultant aggregates. It seems that AMP scenario results in the largest size aggregates, followed by GMP, CMP and UMP, in a descending order.

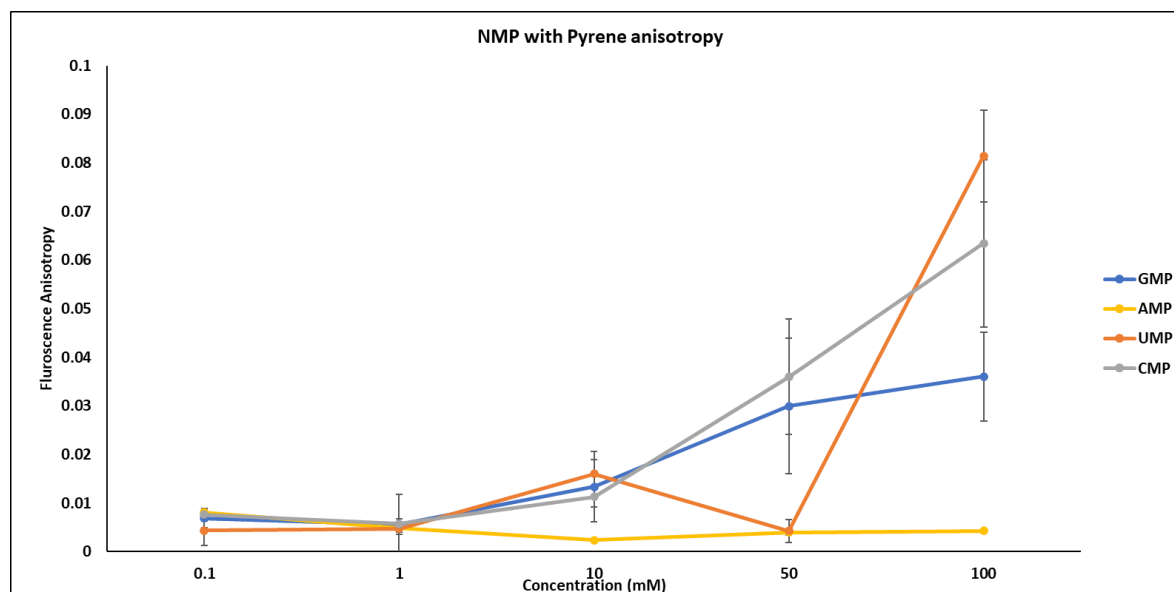


Figure 11: Concentration-dependent change in anisotropy of self-assembled structures of 5' NMPs. The scatter plot shows the anisotropy value of pyrene (Y-axis) with respect to the concentration of corresponding 5' NMPs (X-axis). N=3, error bar=SD.

### Pyrene assay for nucleoside monophosphate in the presence of metal ions

According to the famous J.B.S. Haldane, a proponent of the prebiotic soup hypothesis of life's origins, life on the early Earth came from a soup which is a mixture of different molecules. Given this, the various co-solutes present in this molecular milieu, might have had different effects on self-assembly of nucleotides. Pertinently, metal ions were present in the prebiotic Earth just like in extant biology, wherein they have several important roles for example as cofactors for various enzymes and are also required for cell signaling. So, understanding their effect on prebiotic processes could also help with bridging the huge gap between prebiotic cellular life and extant cellular life. We used sodium, potassium and magnesium for our study. These metal ions are also known to help in the formation of G quartets (Wong et al. 2005). Metal ions were used in different concentrations depending on the metal. Metal ion concentration used were as follows:  $[Na^+] = 100 \text{ mM}$ ,  $[K^+] = 100 \text{ mM}$ ,  $[Mg^{2+}] = 50 \text{ mM}$ . As is clearly evident from Figures 12,13, 14, 15, it can be clearly seen that in all the cases evaluated, the metal ions had a significant effect on the

assembly nature. This data set shows the pyrene spectra for all the 5' NMPs in the presence of metal ions, which is clearly different spectrally from the control (only 5' NMP As observed in Figure 18,  $I_1/I_3$  ratio for all the systems remains similar for all the system. Interestingly, as shown in Figure 19, the excimerization i.e.,  $I_{Ex}/I_1$  we can see that the ratio for GMP goes up after 50mM, for AMP it remains similar, for the pyrimidines there is a decrease in the ratio. These experiments clearly indicate that 5' NMP self-assembly is significantly affected by the co-solutes like metal ions. This experiment clearly demonstrates the role of metal ions on the dynamicity of the aggregates.

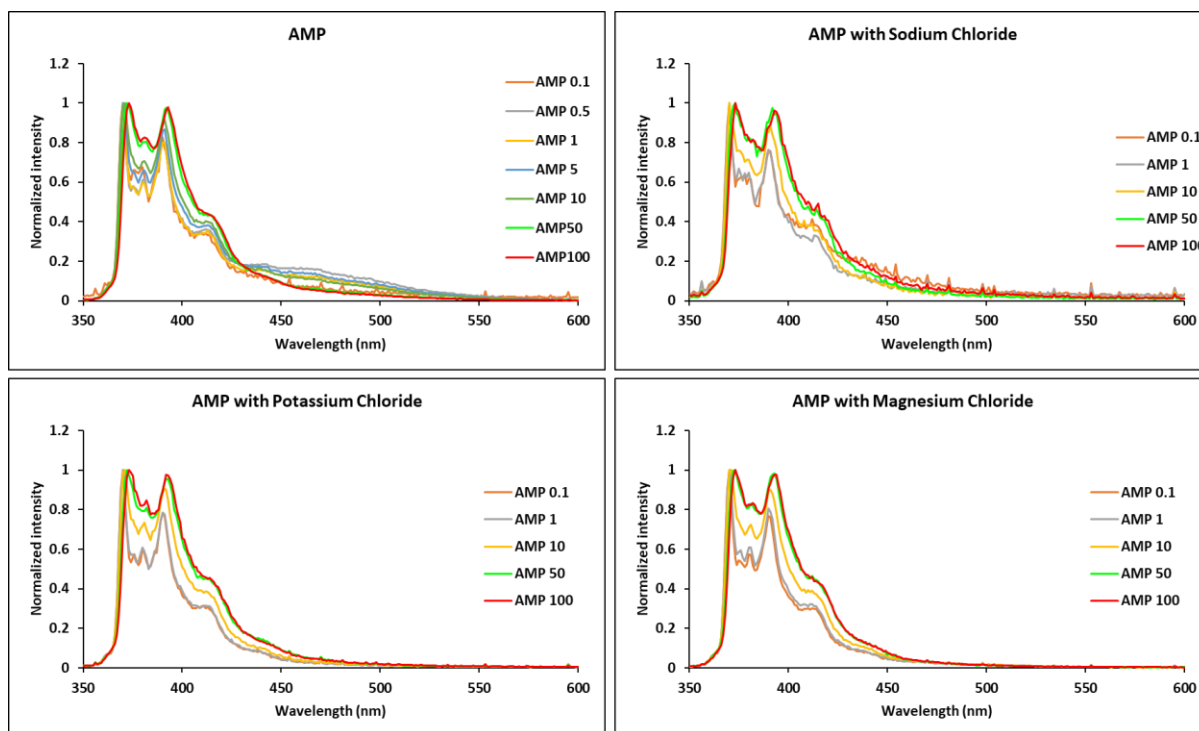


Figure 12: The plots show normalized pyrene emission spectra at different concentrations of AMP solution in the absence and in the presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride). The Y-axis indicates normalized fluorescence intensity and X-axis indicates wavelength (in nm).

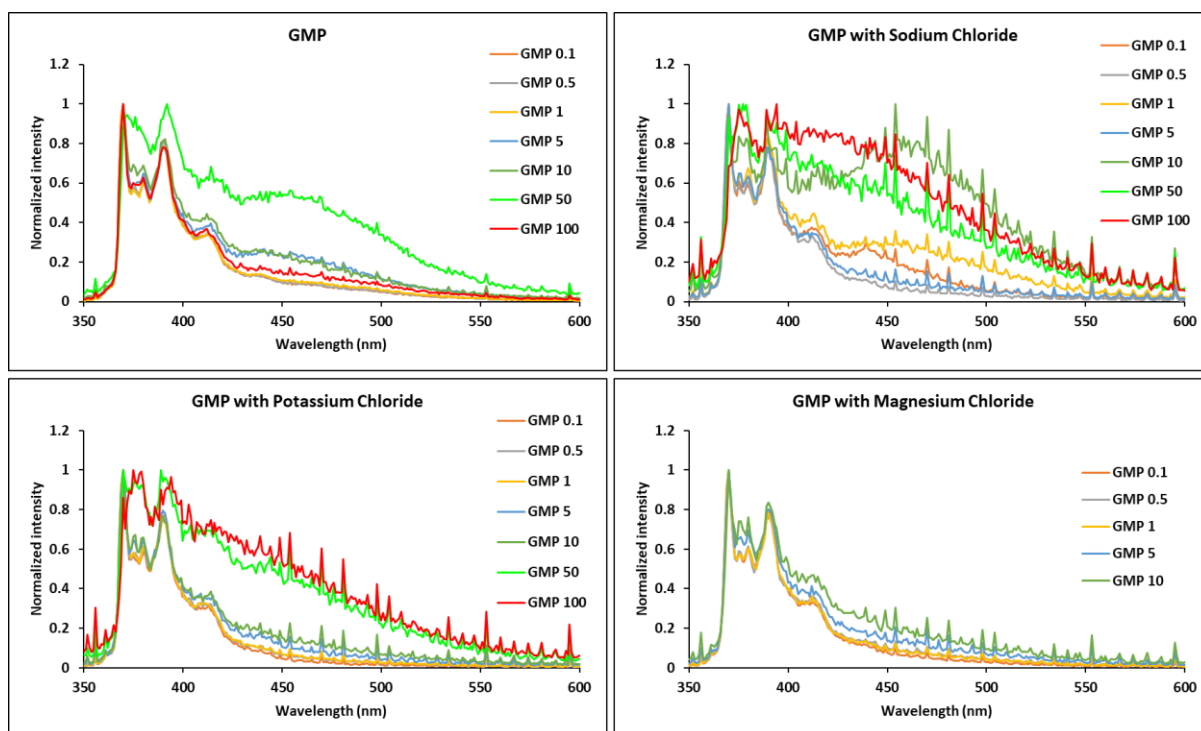


Figure 13: The plots show the normalized pyrene emission spectra at different concentrations of GMP solution in the absence and in presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride). The X-axis indicates normalized fluorescence intensity and X-axis indicates wavelength (in nm).

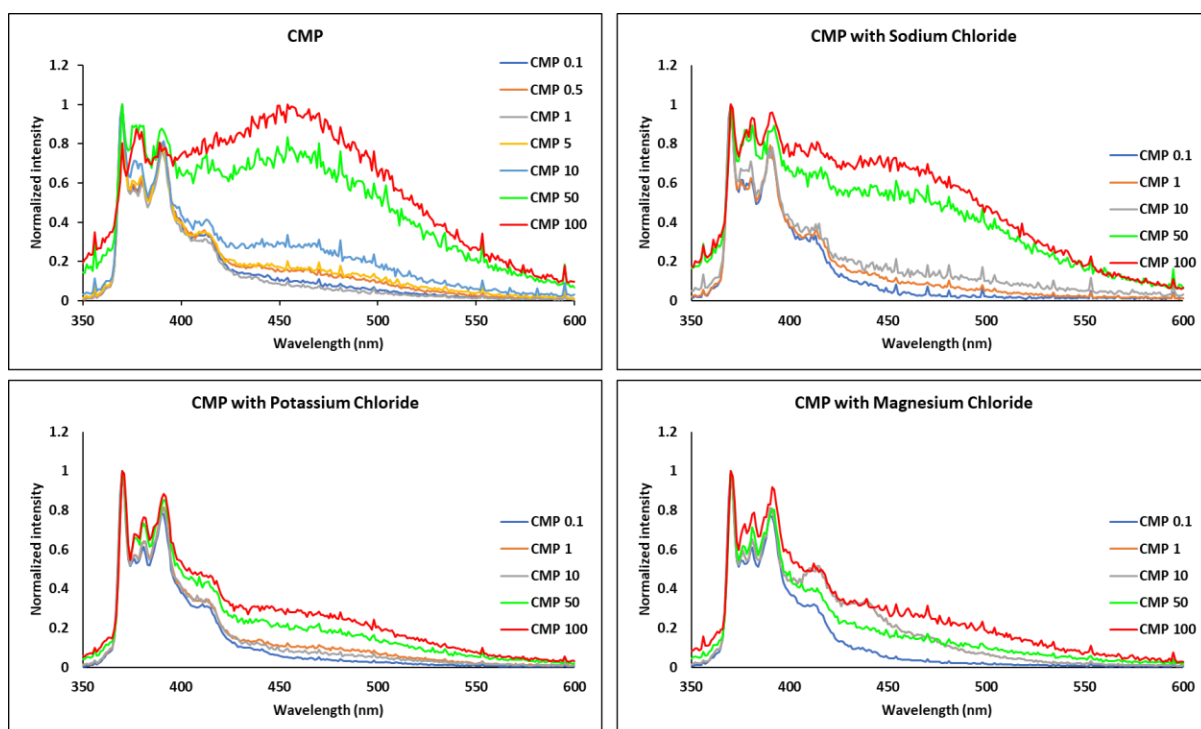


Figure 14: The plots show the normalized pyrene emission spectra at different concentrations of CMP solution in the absence and in presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride). The Y-axis

indicates normalized fluorescence intensity and X-axis indicates wavelength (in nm)

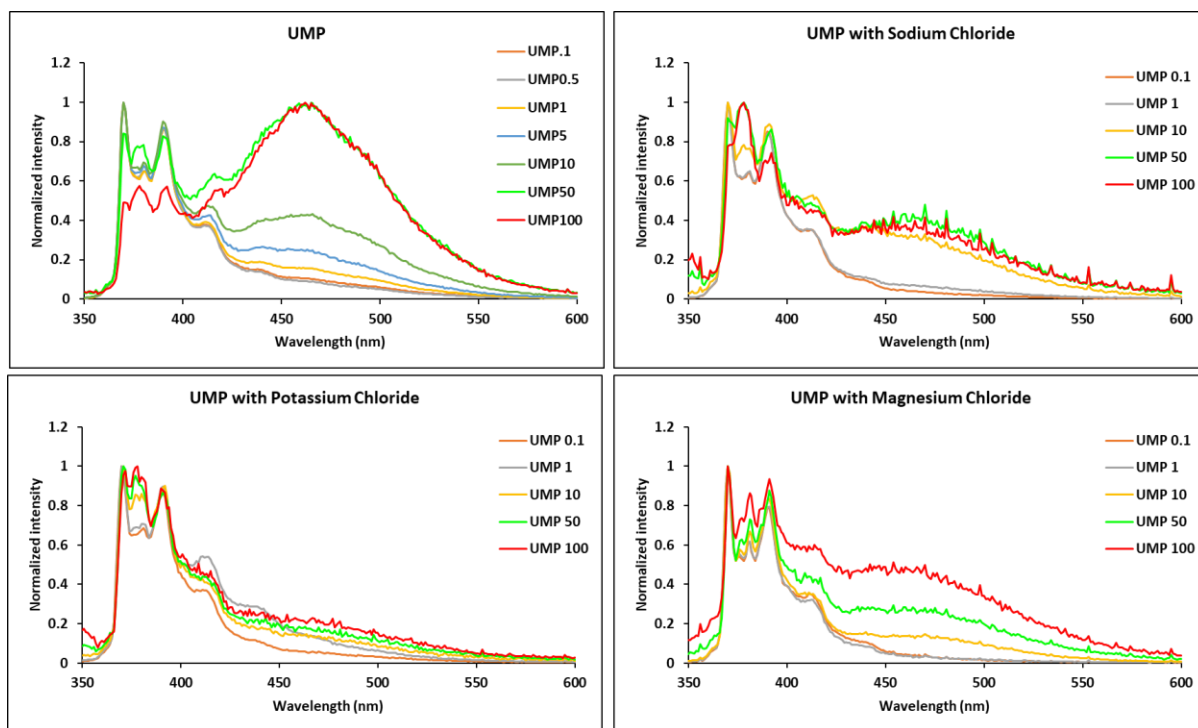


Figure 15: The plots show the normalized pyrene emission spectra at different concentrations of UMP solution in the absence and in presence of various metal salts as indicated (sodium chloride, potassium chloride and magnesium chloride). The Y-axis indicates normalized fluorescence intensity and X-axis indicates wavelength (in nm).

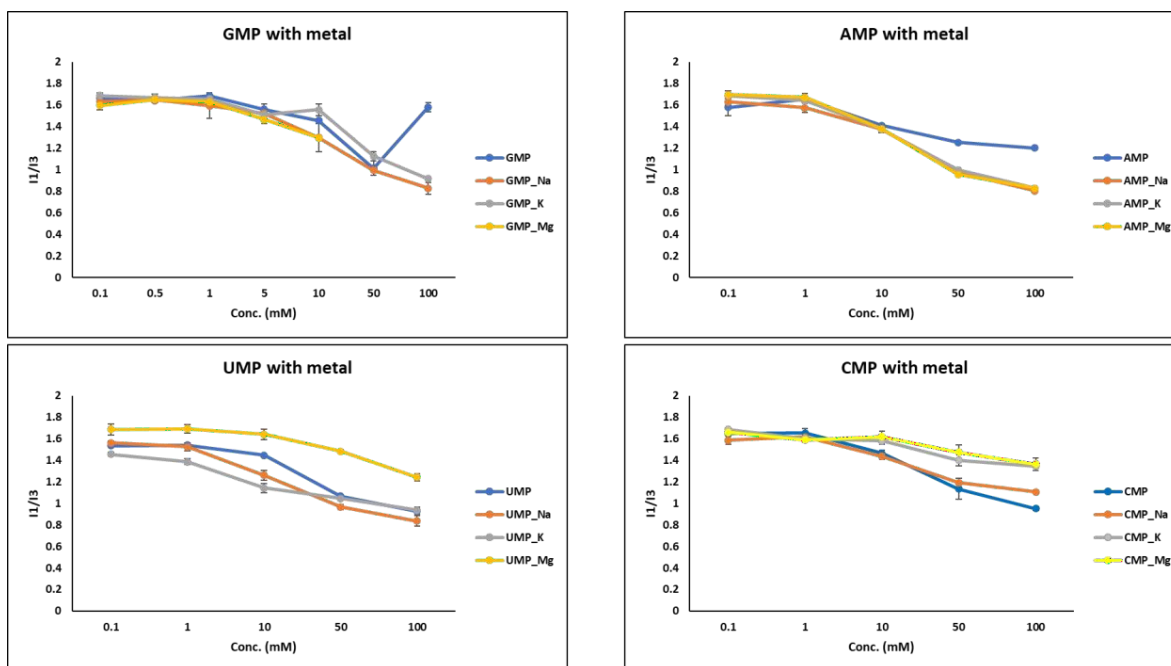


Figure 16: The scatter plot indicates pyrene  $I_1/I_3$  ratios of indicated 5' NMPs in the presence of various metal ions (depicted by different colors). The X-axis indicates the corresponding 5' NMP concentration in mM and the Y-axis indicates the  $I_1/I_3$  ratio. N=3, error bar=SD.

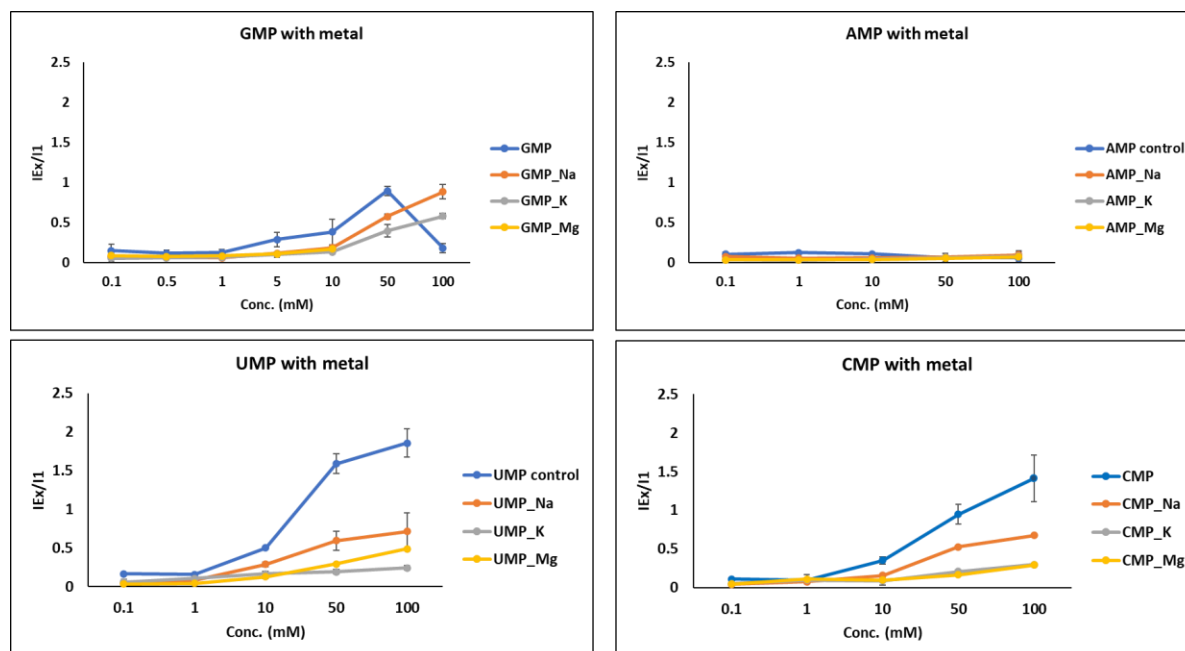


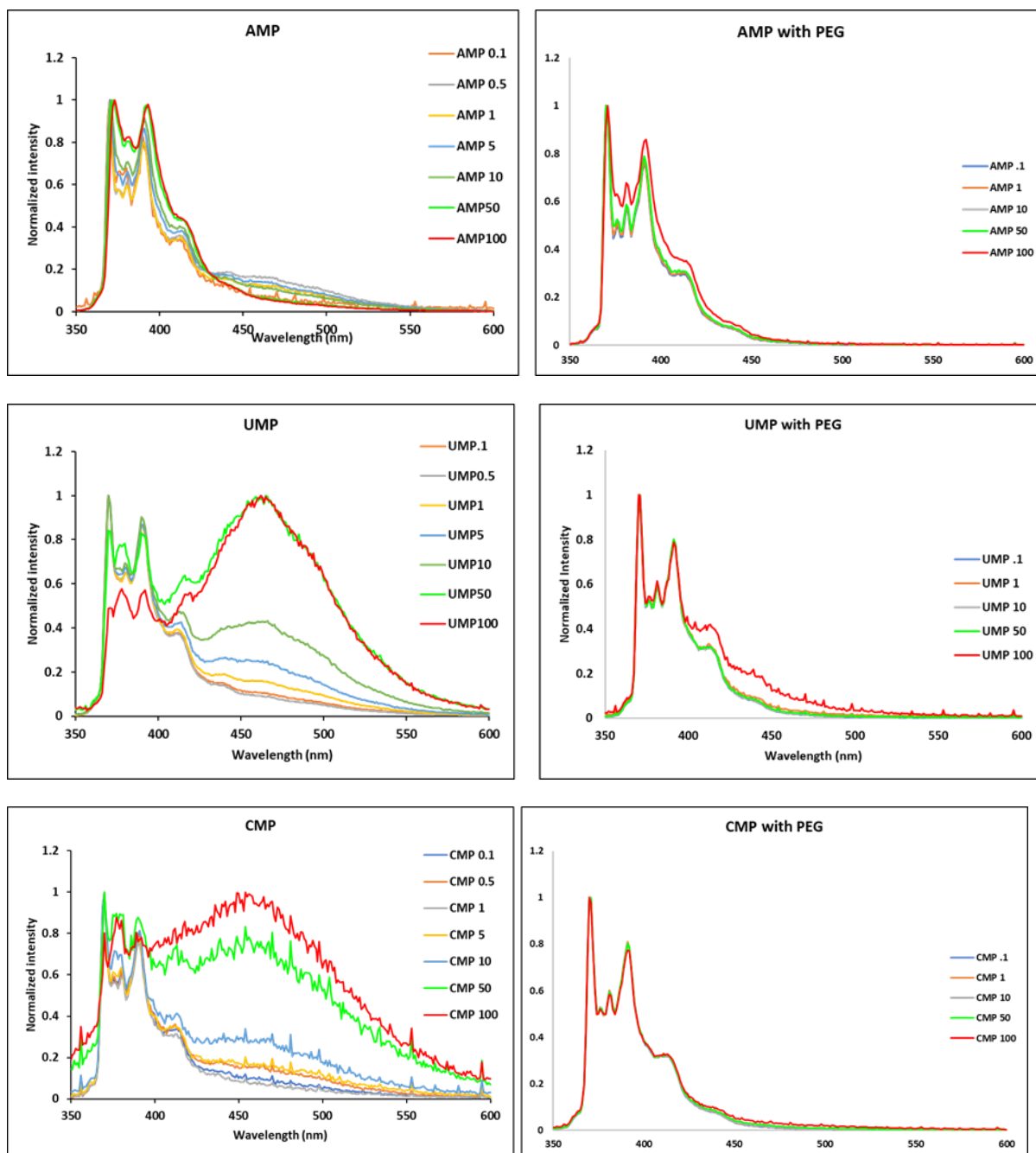
Figure 17: The scatter plot indicates pyrene  $I_{Ex}/I_1$  ratios of indicated 5' NMPs in the presence of various metal ions (depicted by different colors). The X-axis indicates the corresponding 5' NMP concentration in mM and the Y-axis indicates the  $I_{Ex}/I_1$  ratio. N=3, error bar=SD.

## Pyrene assay in the presence of another co-solute proxy, polyethylene glycol (PEG)

As already mentioned, prebiotic soup was a mixture of molecules. In order to mimic this crowding and heterogeneity we used polyethylene glycol (PEG-8000) which is a non-reactive polymer, its major function is to induce a crowding effect which can increase the effective concentration (H. X. Zhou, Rivas, and Minton 2008). PEG used in the reaction was 18% w/v. Previous studies have shown PEG can affect nonenzymatic template directed replication.

In our experiment we added PEG to the sample to make the final concentration 18% w/v. Pyrene was also added to check the dynamicity of the system. Figure 18 indicates pyrene fluorescence spectra for all the 5'NMPs in presence of PEG. Figure 19 shows the quantification of  $I_1/I_3$  ratio. As you can see in the Figure 19 the hydrophobicity goes down for all the NMPs. Compared to control i.e., only 5' NMPs the hydrophobicity is less. From the preliminary results it seems that the PEG is making the aggregates stable because there is very low excimerisation. PEG being a

very large polymer seems to be hindering the formation aggregates. The excimerisation data was not conclusive, further study needs to be done in order to under the role of PEG more clearly.



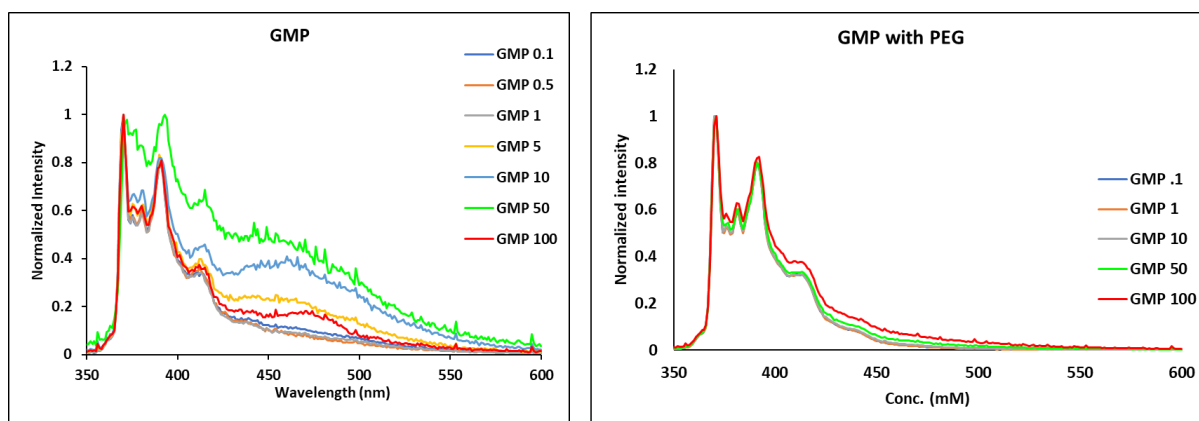


Figure 18: The plots show the normalized pyrene emission spectra at different concentrations of 5' NMP solution in the absence and in presence of 18% w/v PEG. The Y-axis indicate normalized fluorescence intensity and X-axis indicates wavelength (in nm).

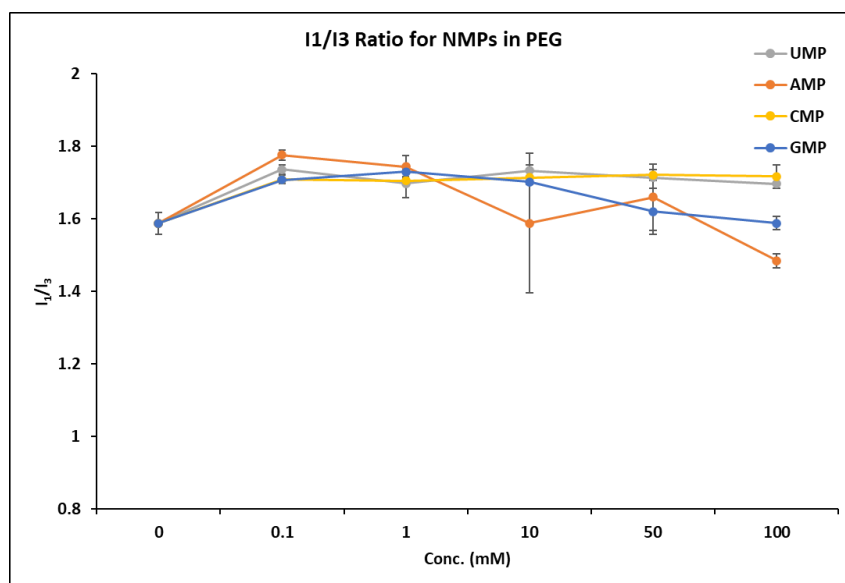


Figure 19: The scatter plot shows pyrene  $I_1/I_3$  ratios of indicated 5' NMPs in the presence of PEG (depicted by different colors). The X-axis indicates the corresponding 5' NMP concentration in mM and the Y-axis indicates the  $I_1/I_3$  ratio. N=2, error bar=SD.

### NMR analysis of self-assembly of nucleoside 5'-monophosphate

Studies of stacking interactions between molecules can be conducted using nuclear magnetic resonance (NMR) spectroscopy. When two aromatic molecules are close enough to one another their pi electron clouds overlap, creating attractive intermolecular forces resulting in stacking interactions. Proton chemical shifts in aromatic rings can be used to gauge the intensity of the stacking interaction using NMR spectroscopy (H. Sigel 2004). Depending on the orientation of the rings and the strength of the connection, when two aromatic molecules stack on top of one



another, the protons in the stacked rings' chemical shifts, may be shifted up field or downfield. The separation between the protons in the stacked rings can influence the size of the chemical shift. In addition to chemical shift measurements, NMR can also be used to study the dynamics of stacking interactions. For example, spin-lattice relaxation times can be used to measure the rate at which molecules exchange between stacked and unstacked conformations. NMR is a powerful tool for studying stacking interactions in solution, providing information on both the strength and dynamics of these important intermolecular forces.

In our experiment we did  $^1\text{H}$  NMR in  $\text{D}_2\text{O}$  to understand the stacking interaction between the nucleotides. We used three different concentrations: 10mM, 50mM, 100mM for all the concentrations. There is an up field chemical shift in the protons of the nucleobase, which shows the stacking interaction. The shift in the proton can be seen in Table 1.

Table 1: The shift in NMR signals for protons of various nucleotides (5' NMPs) at different concentrations in  $\text{D}_2\text{O}$ .

<b>AMP</b>	<b>10mM</b>	<b>50mM</b>	<b>100mM</b>
H16	8.61	8.53	8.46
H10	8.23	8.09	7.97
<b>GMP</b>	<b>10mM</b>	<b>50mM</b>	<b>100mM</b>
H15	8.2	8.16	8.11
H10	5.92	5.88	5.84
<b>CMP</b>	<b>10mM</b>	<b>50mM</b>	<b>100mM</b>
H16	8.13	8.12	8.1
H15	6.01	6.01	6
<b>UMP</b>	<b>10mM</b>	<b>50mM</b>	<b>100mM</b>
H22	8.13	8.11	8.11

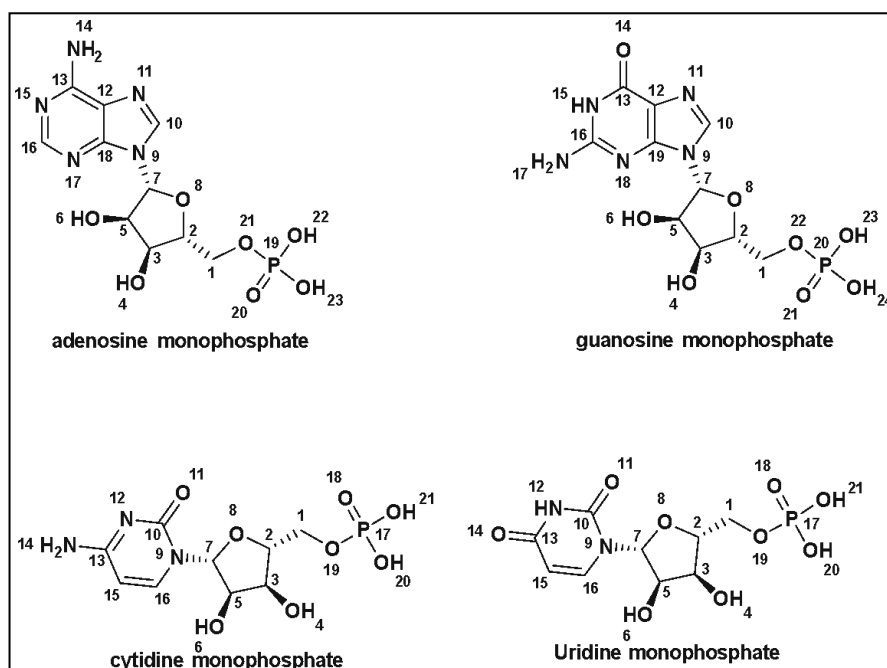


Figure 20: The structures of various 5' NMPs (i.e., AMP, GMP, CMP and UMP) with labelled proton numbers).

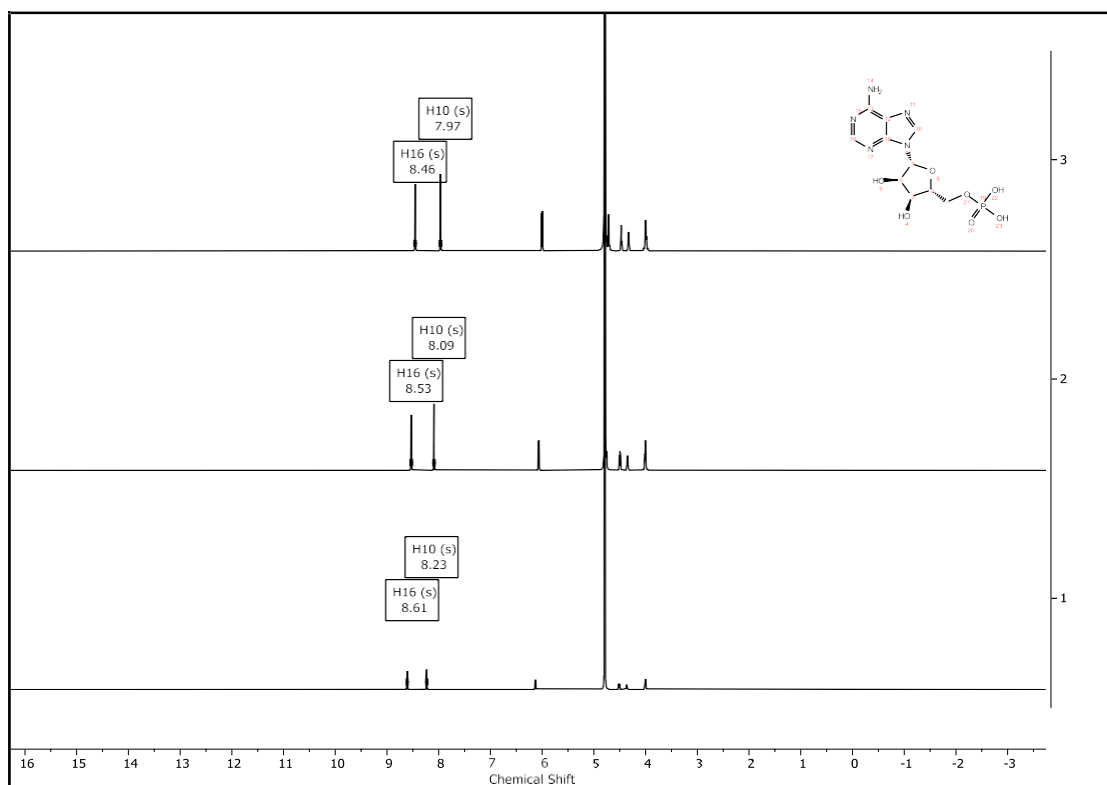


Figure 21:  $^1\text{H}$  NMR characterization of AMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel). Indicated protons i.e., H16 and H10 depict a right shift in these protons showing more shielding potentially due to self-assembled structure formation.

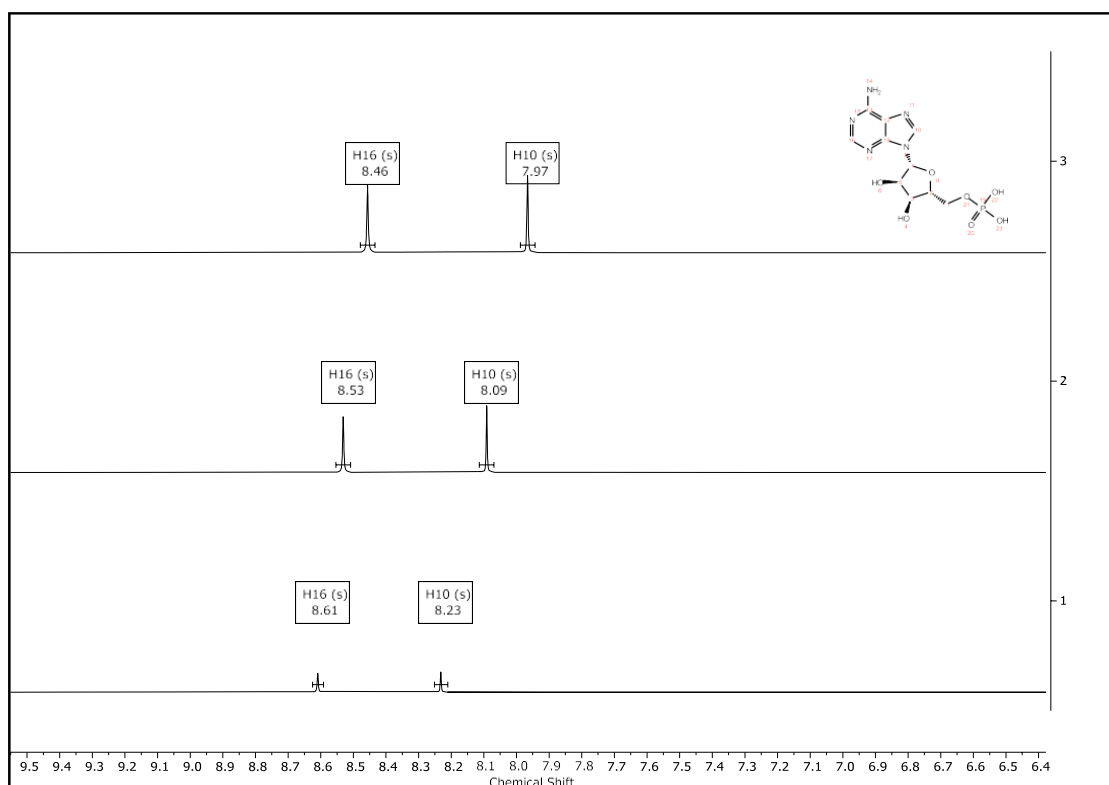


Figure 22: The zoomed version of Figure 21 showing  $^1\text{H}$  NMR spectrum of AMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel). Indicated protons i.e., H16 and H10 depict a right shift in these protons showing more shielding potentially due to self-assembled structure formation.

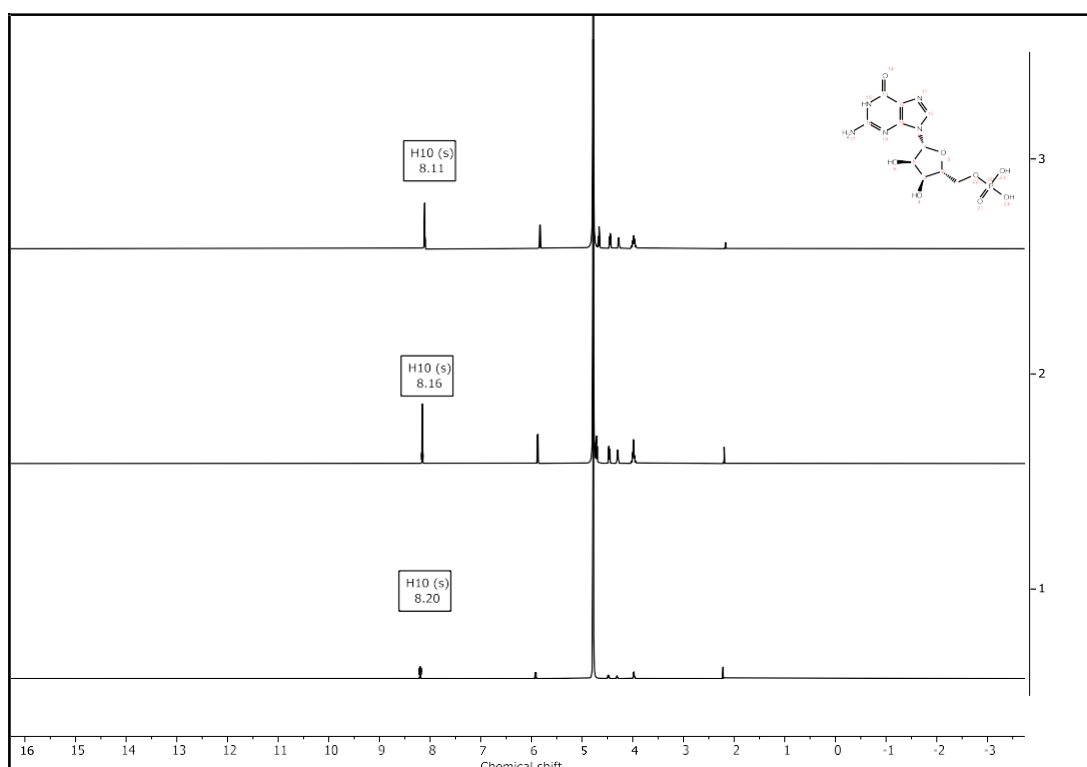


Figure 23: The  $H^1$  NMR spectrum of GMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel). Indicated proton i.e., H10 depicts a right shift in this proton showing more shielding due to self-assembled structure formation.

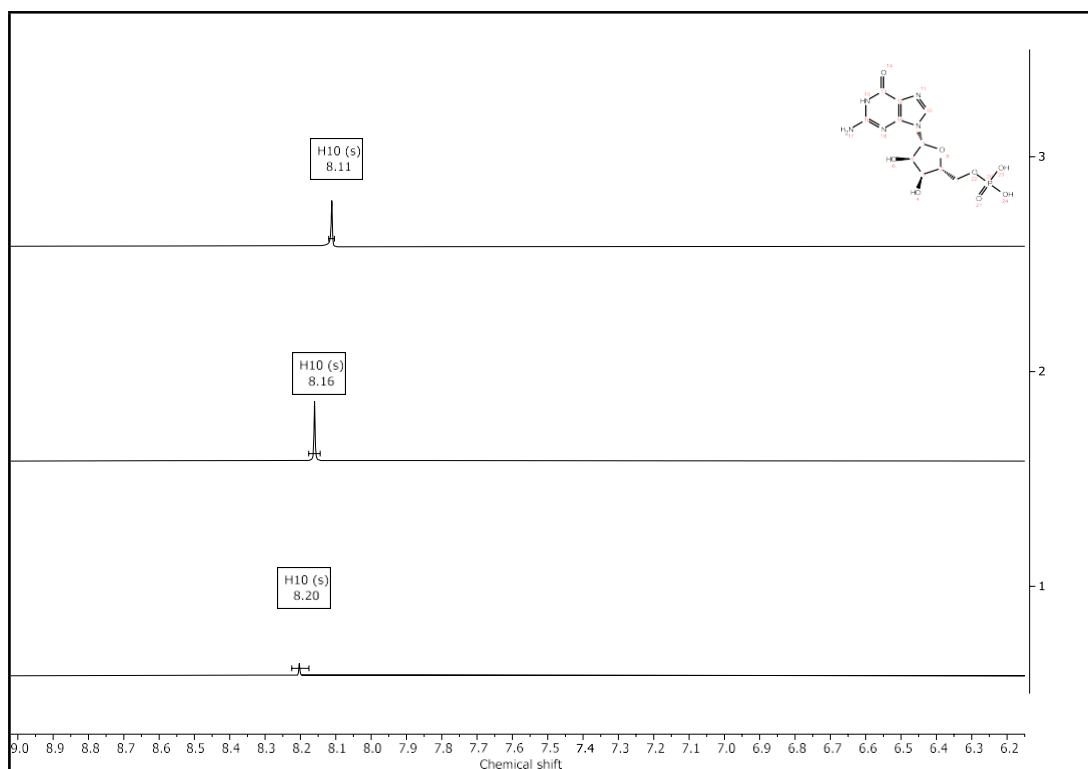


Figure 24: The zoomed version of Figure 23 showing  $H^1$  NMR spectrum of GMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel). Indicated proton i.e., H10 depicts a right shift in this proton showing more shielding due to self-assembled structure formation.

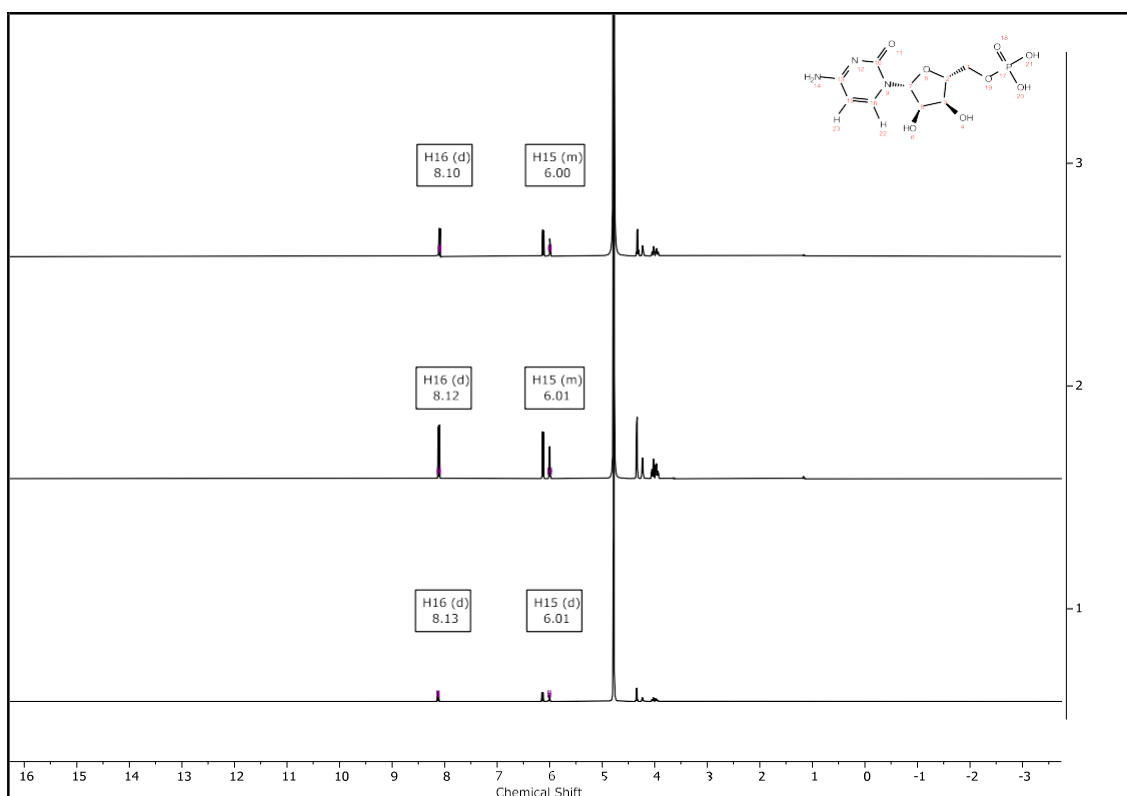


Figure 25: The  $H^1$  NMR spectrum of CMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel). Indicated protons i.e., H16 and H15 depict a negligible shift in these protons showing slight to no shielding change.

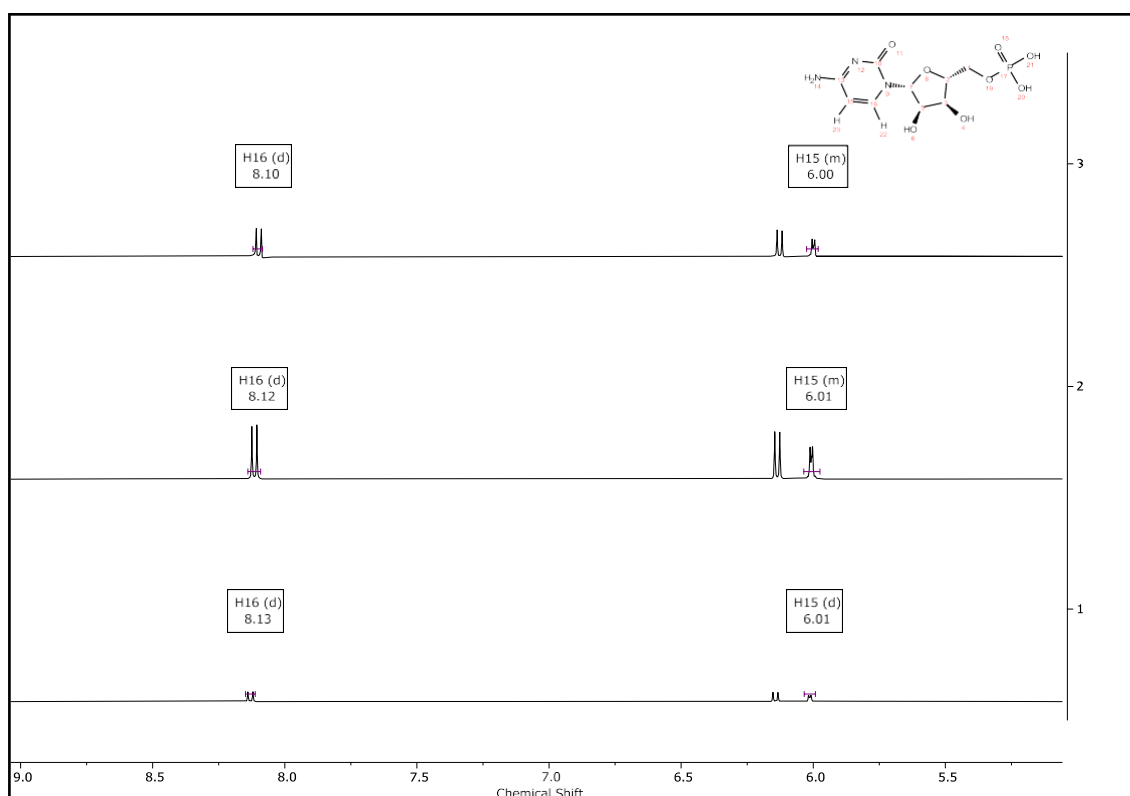


Figure 26: The zoomed version of Figure 25 showing the  $H^1$  NMR spectrum of CMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel).

panel). Indicated protons i.e., H16 and H15 depict a negligible shift in these protons showing slight to no shielding change.

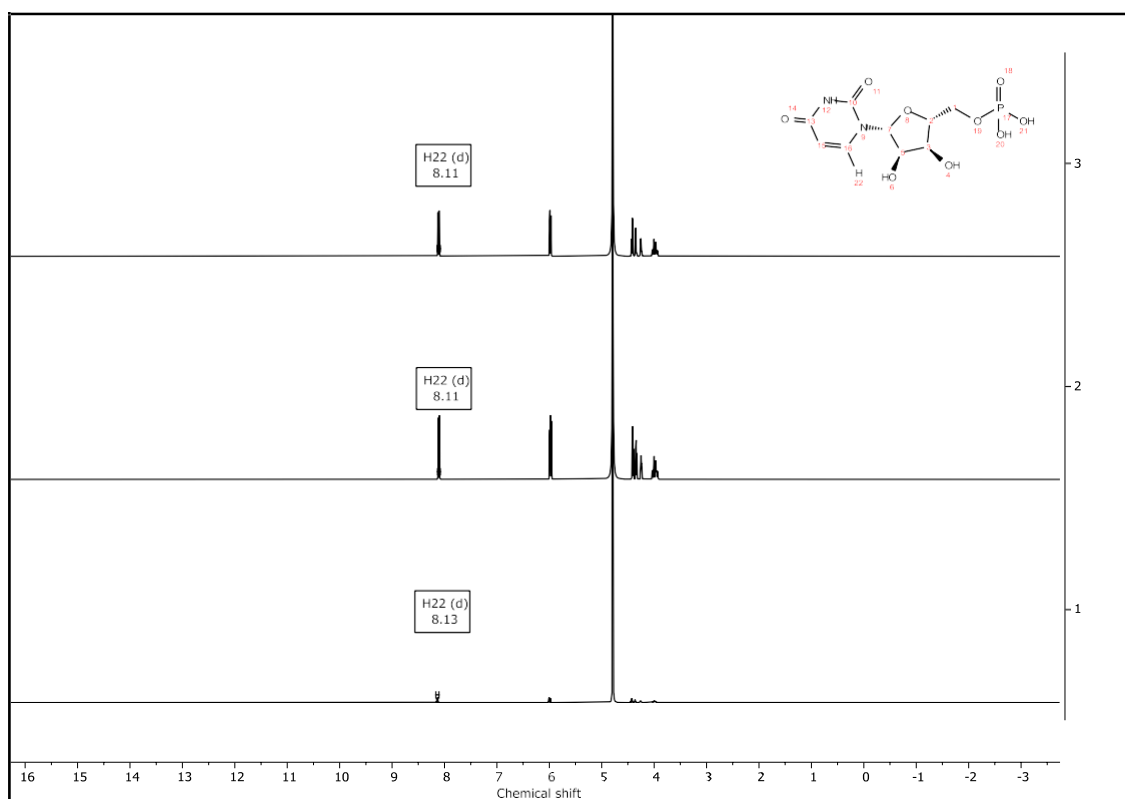


Figure 27: The  $^1\text{H}$  NMR spectrum of UMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel). Indicated proton i.e., H22 depicts a negligible shift in these protons showing slight to no shielding change.

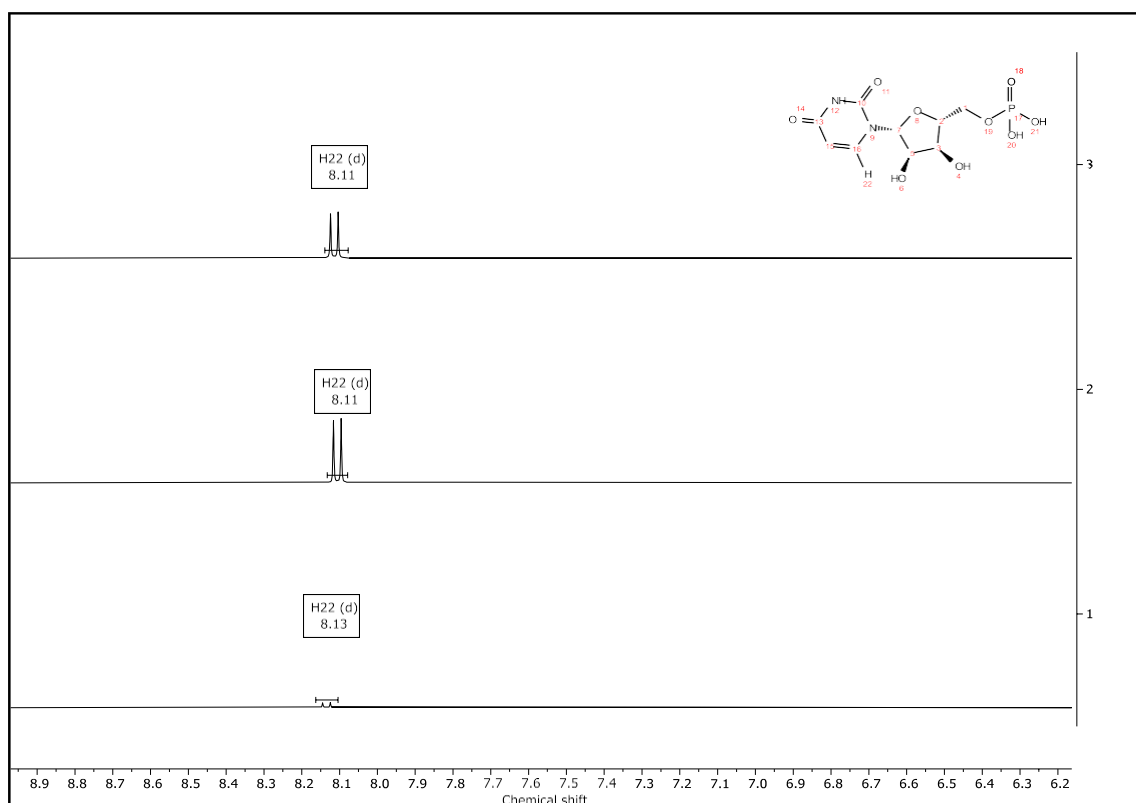


Figure 28: The zoomed version of Figure 27 showing the  $H^1$  NMR spectrum of UMP at different concentrations: 10 mM (lower panel), 50 mM (middle panel) and 100 mM (upper panel). Indicated proton i.e., H22 depicts a negligible shift in these protons showing slight to no shielding change.

## 2D NMR NOESY for Understanding the interactions

NOESY is a technique used for studying the spatial relationships between protons in a molecule. It is used to determine the relative positions and distances of protons in a molecule, and can be used to investigate the structure, dynamics, and interactions of molecules in solution. NOESY spectra is a 2D spectra in which the chemical shift of the protons in a NOESY experiment is represented by the first dimension of the spectrum, and the transfer of magnetization from one proton to another is represented by the second dimension. The nuclear Overhauser effect (NOE), a phenomenon of quantum mechanics, resulting from the interaction of nuclear spins, is responsible for the transfer of magnetization. The NOE happens when a proton's magnetic field interacts with another proton's resonance frequency. A cross peak in the NOESY spectrum indicates that there has been a transfer of magnetization between two protons when they are close to one another in space as a result of the NOE.

In order to study stacking interactions, NOESY can provide information on the orientation of two adjacent nucleotides or aromatic rings in a molecule, which can help to determine the type of stacking interaction present. In our experiments we can clearly see the correlations between different protons for all the 5' NMPs.

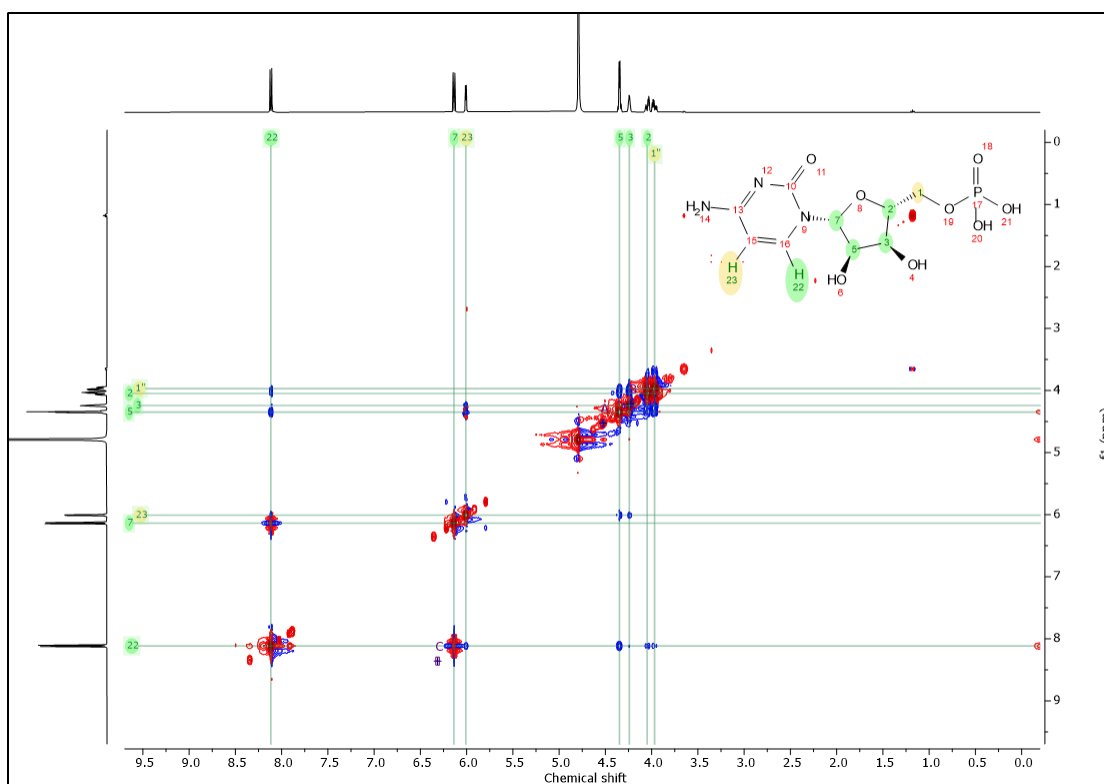


Figure 29: This Figure shows 2D NOESY spectrum for CMP at 50mM concentration recorded in D<sub>2</sub>O. H22-----H23, H22----H7, H23 --- H5 are the proton interactions that are clearly visible in this spectrum.



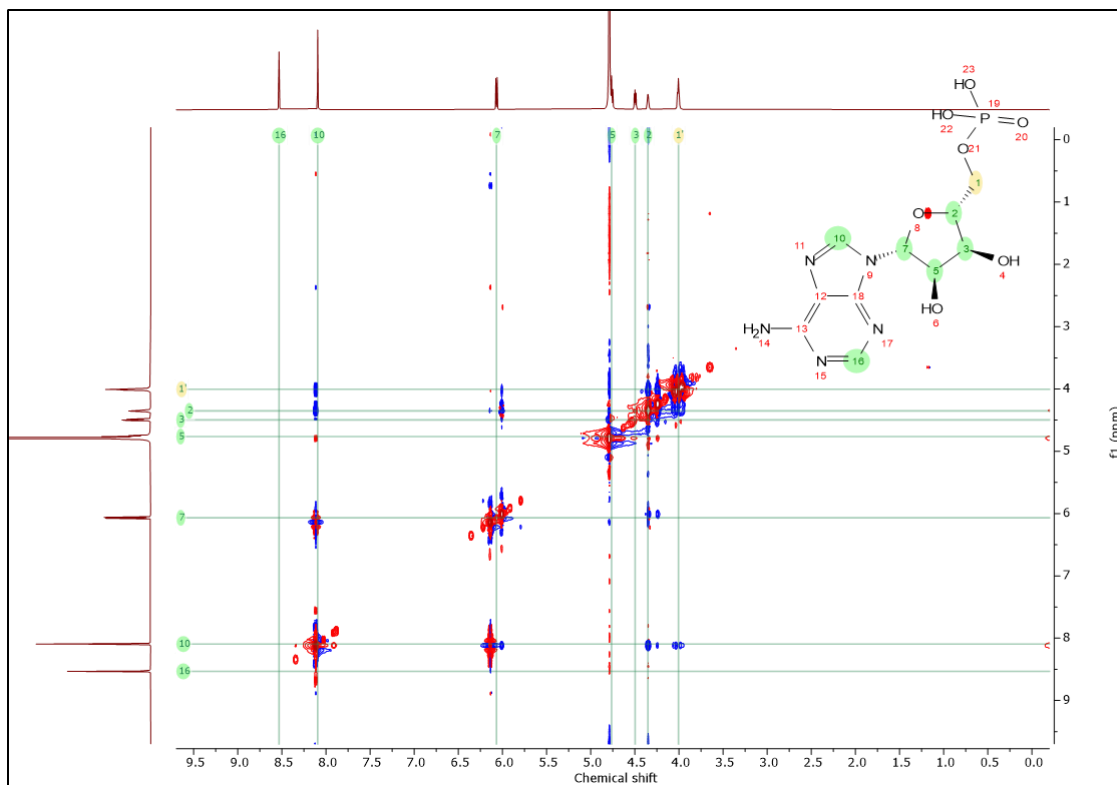


Figure 30: This Figure shows 2D NOESY spectrum for AMP at 50mM concentration recorded in D<sub>2</sub>O. H10 ↔ H7, H10 ↔ H2, H7 ↔ H2 are the proton interactions clearly visible in this spectrum.

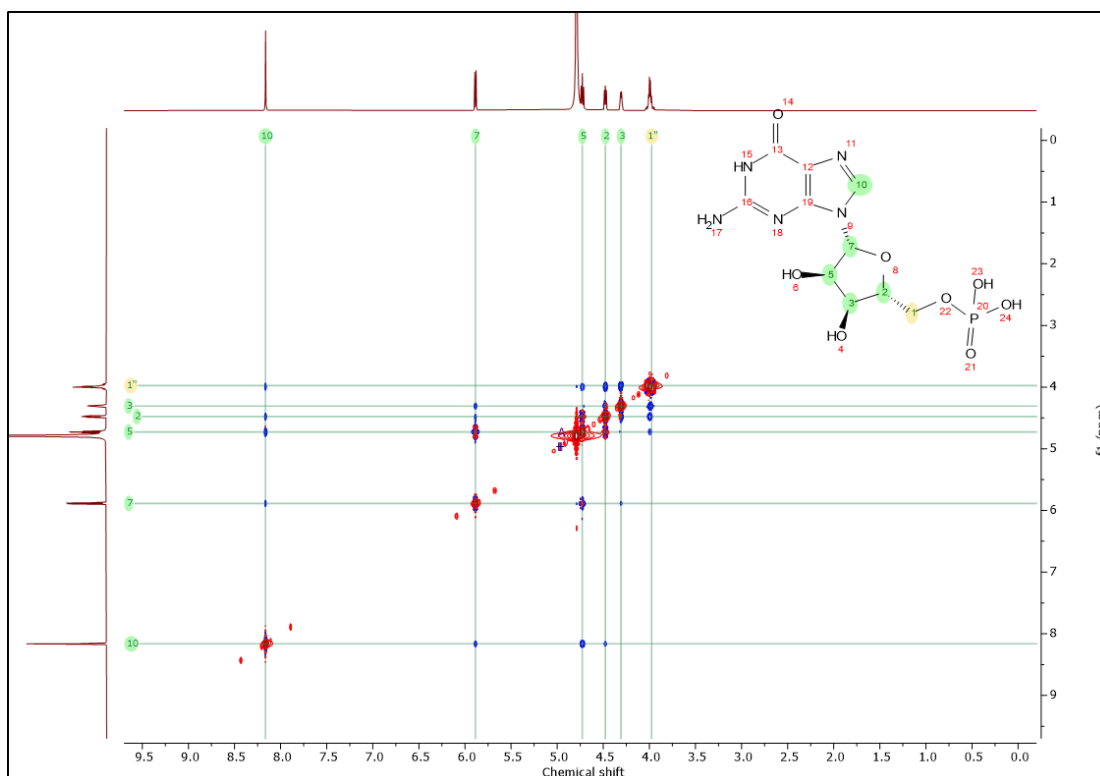


Figure 31: This Figure shows 2D NOESY spectrum for GMP at 50mM concentration recorded in D<sub>2</sub>O. H7----H10, H10----H5, H5---- H7 are the proton interactions visible in this spectrum.

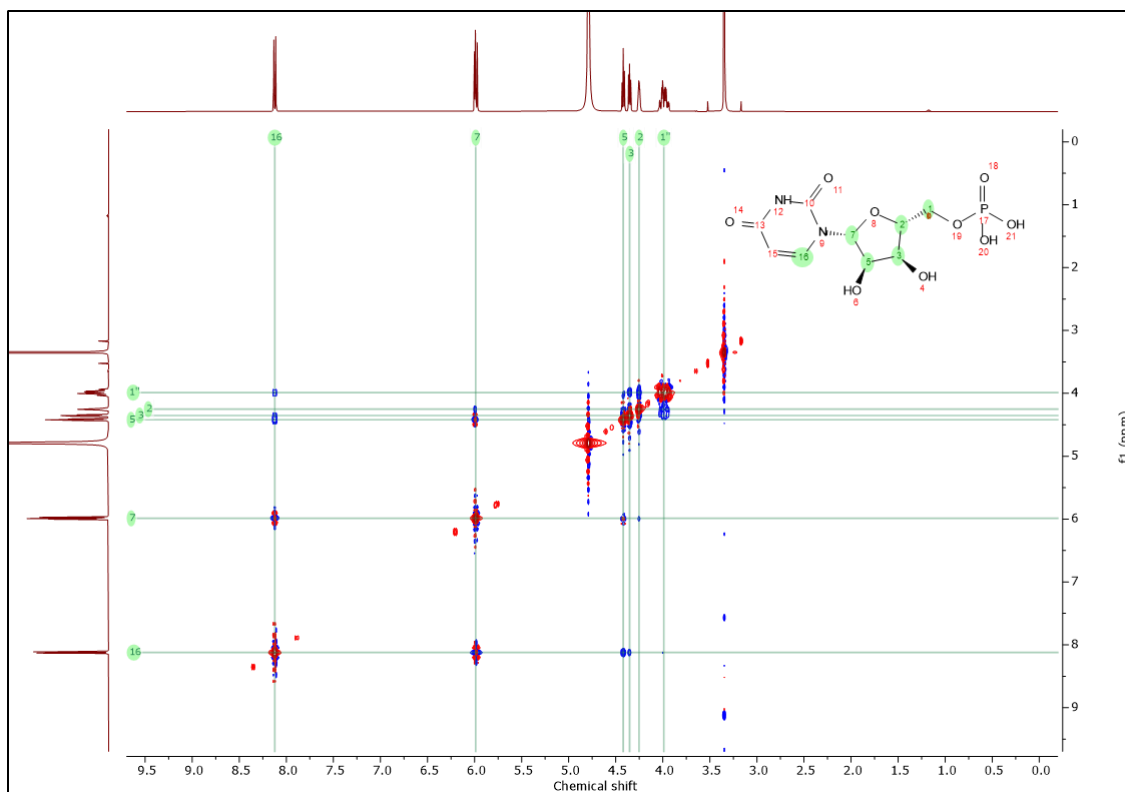


Figure 32: This Figure shows 2D NOESY spectrum for UMP at 50mM concentration recorded in D<sub>2</sub>O. H16----H5, H16----H3, H16----H7 are the proton interactions visible in this spectrum.

In Figure 29 ,30 ,31 ,32 we can see the cross peaks of the 2D spectra that shows which protons can interact with which other protons. All the experiments were done in D<sub>2</sub>O in 400 MHz NMR spectrometer. Stronger interaction has an increase in signal intensity. These interactions help us to determine the spatial arrangement of the protons and thus the molecules. From the above spectra it can be clearly seen that there is spatial interaction between the nucleotides, which strongly supports the formation of higher order aggregates at the respective 5' NMP concentrations.

## **Chapter 3: Implications of self-assembly on prebiotically relevant processes**

It is believed that nucleotides, which were present on the early Earth, served as the fundamental components for RNA and DNA. These nucleotides were thought to have formed through non-enzymatic processes driven by specific environmental conditions. In the origins of life field, nucleotides are primarily viewed as the precursors to genetic material. However, due to their ability to engage in various interactions, they can also create diverse complexes that might have had different functions.

### **Nucleotides as catalyst**

Nucleotide self-assembly can have varying implications on early Earth. As mentioned above, nucleotides have a variety of functional groups due to which it can show multiple interactions with different molecules such as metal ions, minerals, lipids and other prebiotically relevant co-solutes. With metal ions, nucleotides can interact either via electrostatic interaction like in the case of  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{+2}$ , or via forming coordination bonds with transition metal ions such as  $\text{Cu}^{+2}$ ,  $\text{Co}^{+2}$  etc., to result in coordination complexes (Lopez and Liu 2017). Inspired by this, and given the plausibility of such complexes to form readily, along with their enhanced stability as compared to native enzyme, studies have investigated the capability of these complexes to mimic the activity of native enzymes in order to make synthetic/artificial enzymes (Wong et al. 2005).

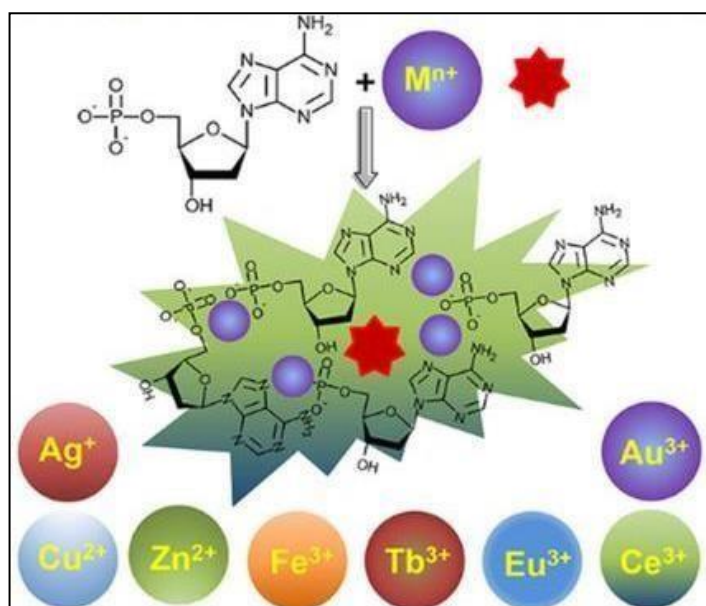


Figure 33: The illustration depicts how the presence of metal ions can induce the self-assembly of 5' NMP (this is showing GMP as an example in this context). This image was adapted from (Lopez and Liu 2017)

Towards this, GMP has been well-explored due to its tendency to form G-quartets/hexads via Hoogsteen base pairing and are stabilized by metal ions such as: by  $\text{Na}^+$  and  $\text{K}^+$ . In the case of transition metal ions, theory suggests that the coordination happens via the 'O' atom of phosphate group and the 'N' of the nucleobase to form different complexes (Griesser et al., n.d.). These coordination complexes have been demonstrated to mimic enzymes. One of the studies has shown that GMP, in presence of  $\text{Cu}^{2+}$  can form metal organic frameworks (MOFs), which has the capability to catalyze redox reaction. These MOFs can mimic laccase enzymes, which can perform redox reactions. The MOF forms were quite stable over a large pH range and had similar or even more catalytic activity than naturally occurring laccase enzymes (Liang et al. 2017). This gives us sufficient basis to posit that such MOFs can act as catalysts and can be readily formed in a prebiotic setting. These can, therefore, be thought of as the simplest form of metabolic molecules, which can perform reactions. Nonetheless more work needs to be done in this area to better understand how and what kinds of prebiotic reactions they can catalyze.

### **Nucleotides in solubilizing hydrophobic molecules**

As already mentioned, nucleotides can indulge in multiple interactions. One of the interactions it participates in is pi-pi stacking, in which the aromatic system of the nucleobases can interact with other aromatic systems. Previous studies have shown

interaction with aromatic systems like bipyridine, phenanthroline etc., by calculating the association constants (H. Sigel 2004). In the presence of metal ions like  $\text{Cu}^{2+}$ , the association constant increases. Hence, it is fair to think that prebiotic molecules having aromatic systems could have interacted with amino acids like tyrosine, phenylalanine etc., with potentially interesting implications for prebiotic chemistry.

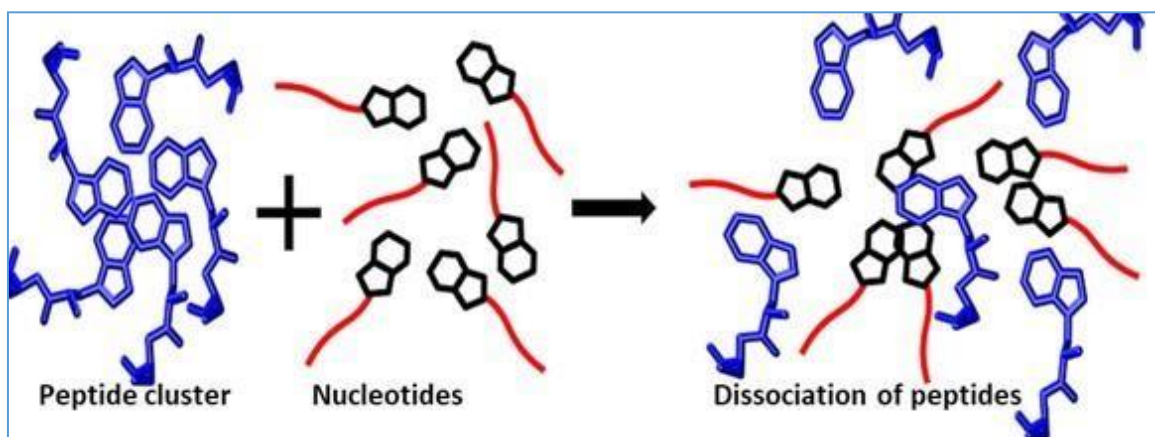


Figure 34: The illustration depicts how nucleotides could have driven the solubilization of hydrophobic molecules (peptide clusters are shown here as an example). This image reproduced from (Pandey et al. 2022)

In extant biology we see a variety of amino acids, ranging from the polar to the aromatic residues. The low solubility of aromatic amino acids in water would have hindered the addition of these amino acids in any prebiotic peptide formation on the early Earth. For example, let's imagine there is a peptides formation reaction going on in hydrothermal vents on the prebiotic Earth. The aromatic amino acids are poorly soluble in water, hence there will be very low to minimal aromatic amino acid concentration available for participating in the peptide formation reaction. However, the solubilization property of nucleotides could have helped in dissolving these aromatic amino acids in water, thus making them available for such reactions. This is one of the ways by which nucleotides could have facilitated prebiotic reactions.

In order to understand this we used Fluoresceine diacetate (FDA) which can act as a proxy for the hydrophobic molecules in early Earth. We check the solubility of FDA in presence of 5`NMPs which showed increase as we increased the concentration of nucleotides. We also wanted to understand its effect on nonenzymatic replication we used bipyridine (BPY) which acted as a proxy for all the aromatic rings. Addition of BPY in the reaction reduced the rate and in order to understand the interaction we used UV/vis spectroscopy.

## **Material & Methods**

All materials used in the subsequent experiments were reagent grade and used without further purification. Disodium salts of guanosine 5'-monophosphate, cytosine 5'-monophosphate, adenosine 5'-monophosphate, uridine 5'-monophosphate were obtained from Sigma Aldrich. 2,2 Bipyridine was obtained from Thomas Baker.

### **UV/Vis spectroscopy**

The absorption spectra were recorded with a UV/Vis spectrophotometer where the absorbance reading was taken from 200-600nm. The main stock of 5' NMPs was prepared in filtered Milli-Q water, while 2,2 bipyridine was prepared in DMSO and further dilutions were made in Milli-Q water. UV/Vis spectrometry was done by preparing samples of each concentration and then placing them into instrument and took their readings.

### **Urea PAGE analysis for nonenzymatic primer extension**

To investigate non-enzymatic template-directed replication, a urea polyacrylamide gel electrophoresis (PAGE) method was utilized. Time points were collected in TBE buffer containing 8 M urea and 100 mM EDTA. The samples were heated at 90 °C along with an excess amount of primer (ten times more than the labelled primer) to separate the fluorescently labelled (Cy3) reaction primer from the template. Subsequently, the time points were subjected to a 20% denaturing PAGE to determine the presence of both the fluorescently labelled primer and the extended product. The gels were then scanned using a Typhoon Trio plus imager (GE Healthcare), and the scans were analysed using ImageQuant v 5.2 software. The extension of the primer over time was calculated by dividing the intensity of the unreacted primer by the sum of the intensities of both the reacted (extended) and unreacted primer. A linear equation was fitted to multiple early data points to estimate the initial reaction rate.

## **Results:**

### **Solubilization of hydrophobic molecules**

The prebiotic soup had a diverse set of molecules, ranging from small molecules to large polymers. Some of those molecules were hydrophobic in nature so their solubility would have limited. Nucleotides, as already mentioned, can show many interactions and because of pi-pi interaction it can solubilize aromatic hydrophobic molecules ranging from amino acid to PAHs etc.

In this experiment we used fluorescein diacetate (FDA), which is hydrophobic dye. Because of its low solubility in water, it remains in aggregated form and shows low absorbance. When we add nucleotides in the system, FDA disaggregates, increasing its dissolution in water, which results in an increase in its absorbance value. The absorbance was taken at 498nm.

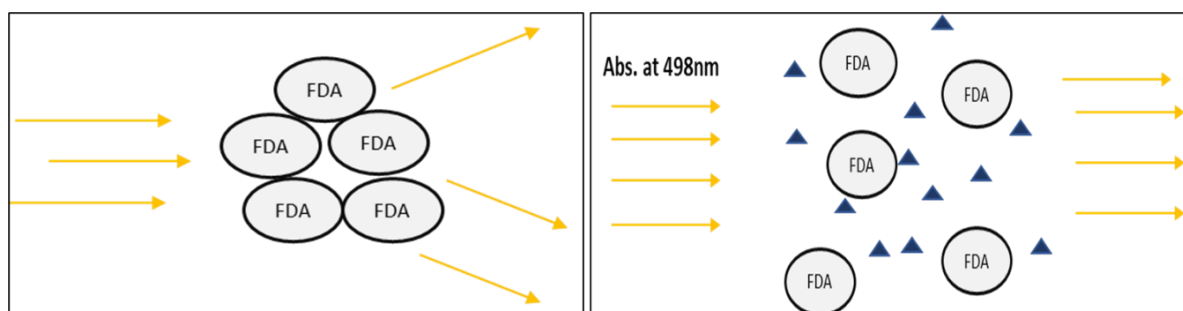
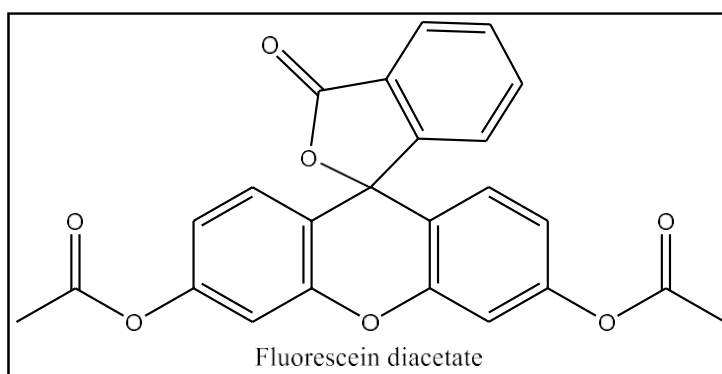


Figure 35: Schematics showing FDA absorption assay to check for the solubilizing capabilities of 5' NMPs.

As you can see in Fig 36, the purines i.e., AMP and GMP, show a large increase in absorbance value when compared to the pyrimidines i.e., UMP and CMP. This is not fully surprising as pyrimidines are known to have lower stacking capabilities than purines, and hence are not able to dissolve the FDA as readily as purines. This experiment shows that the nucleotides can solubilize the FDA dye and it also shows the comparison of solubilization capacity of different nucleotides.

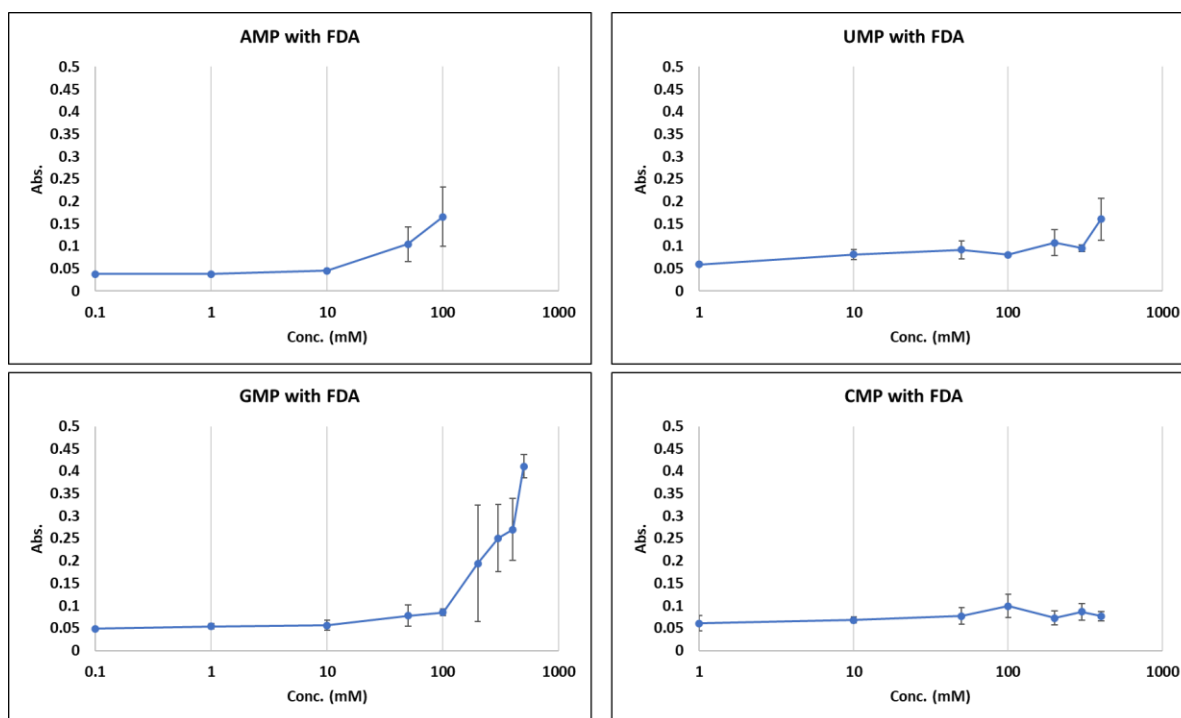


Figure 36: The solubilization of FDA in the presence of different 5' NMPs at various concentrations. The Y-axis shows absorbance at the  $\lambda_{\max}$  of FDA i.e., 498 nm and the X-axis depicts various concentrations of corresponding 5' NMPs (in mM). An increase in the absorbance value was observed for all the 5' NMPs indicating solubilization of FDA with increasing concentration of 5' NMPs. N=3, Error bars = S.D.

### Nonenzymatic template directed replication

On the prebiotic Earth, the formation of enzymes would have been non-trivial; so, in order for RNA to replicate it would have had to be enzyme free. Non-enzymatic replication of RNA may serve as an intermediate step between the prebiotic synthesis of nucleotides and the RNA world characterized by the presence of RNA enzymes (ribozymes) that may have facilitated replication in ancient cells containing RNA genomes. In the lab, we use 5' activated nucleotides (imidazole activated versions called imidazolides), which are incorporated against the respective templating base in the templates, to study nonenzymatic replication reactions. The activated nucleotides have a good leaving group, which makes it attack the nucleophilic  $\text{NH}_2$  center on the 3'- $\text{NH}_2$  terminated primer, leading to an extended primer product. This reaction gives us a plausible mechanism on how RNA could have replicated under prebiotic conditions. Importantly, this framework also allows us to study how different prebiotically relevant co-solutes, could have affected these reaction rates.



For nonenzymatic template-directed replication, the RNA primer used was Primer Amino G, which terminated with 3'-amino-2', 3'-dideoxynucleotide (Metkinen, Finland) that was acquired from Keck laboratory, Yale, USA. The primers are labelled with Cy3 on the 5' end, which allows for its detection after getting resolved on polyacrylamide gel electrophoresis (PAGE) analysis. The RNA templates used in the reaction were acquired from Thermos Scientific (Dharmacon). Figure 37 shows the sequences used for the primer and templates, where the red color indicates the base of the template.



Figure 37: Primer sequence and template sequence used in the nonenzymatic template replication. Template has highlighted base to which incoming nucleotide base pairs.

In order to facilitate the efficient binding between the template and primer before starting the reaction, a combination of 0.325  $\mu$ M primer and 1.3  $\mu$ M template was heated at 90°C for 5 minutes and then cooled to room temperature for another 5 minutes. Tris (100mM pH 7.0) and NaCl (200 mM) were also added to the reaction mixture. Adenosine 5'-phosphoimidazole (ImpA) was introduced to initiate the primer extension reactions with a final concentration of 10 mM. To better understand the impact of other solutes on the reaction rate, we performed a test reaction by adding cognate nucleotide ImpA across its template base U in the presence of BPY as a co-solute, while a control reaction was conducted without BPY. Refer to Table 2 for the complete list of reaction components.

Table 2: Summary of reactants used and their concentrations in a typical template-directed replication reaction. Test has BPY and control one has no BPY in the reaction mixture.

Chemical	Test	Control
Template (1.3 $\mu$ M)	+	+
Primer (0.325 $\mu$ M)	+	+
NaCl (200 mM)	+	+
Tris buffer (100 mM)	+	+
ImpA (10 mM)	+	+
BPY (10 mM)	+	-

Bipyridine is an aromatic hydrocarbon which can act as a proxy for different aromatic rings. BPY has also been shown to have a tendency to stack with nucleotides. Therefore, we used urea-PAGE analysis to calculate the rate of nonenzymatic template-directed replication, both in the presence and absence of BPY (as detailed in the methods section). The rate of addition of the monomer to the primer-template complex was significantly decreased in the presence of BPY. This may be mainly due to the interaction of BPY and 5' NMPs that somehow affected the availability of the ImpA for the extension reaction.

We found that the reaction rate in the presence of BPY was 0.264/hr, while that for control was 0.477/hr. We saw a decrease in the rate of reaction by around 44.6% when BPY was added in the system. This result suggests that there is some interaction that could be happening because of which the activated nucleotides are not readily available for primer extension, thereby leading to reduction of the extension rate. Is it because the nucleotides were forming a complex with BPY possibly via Pi-Pi stacking. This possible complex formation between nucleotides and BPY was investigated in the next experiment using UV/Vis spectroscopy.

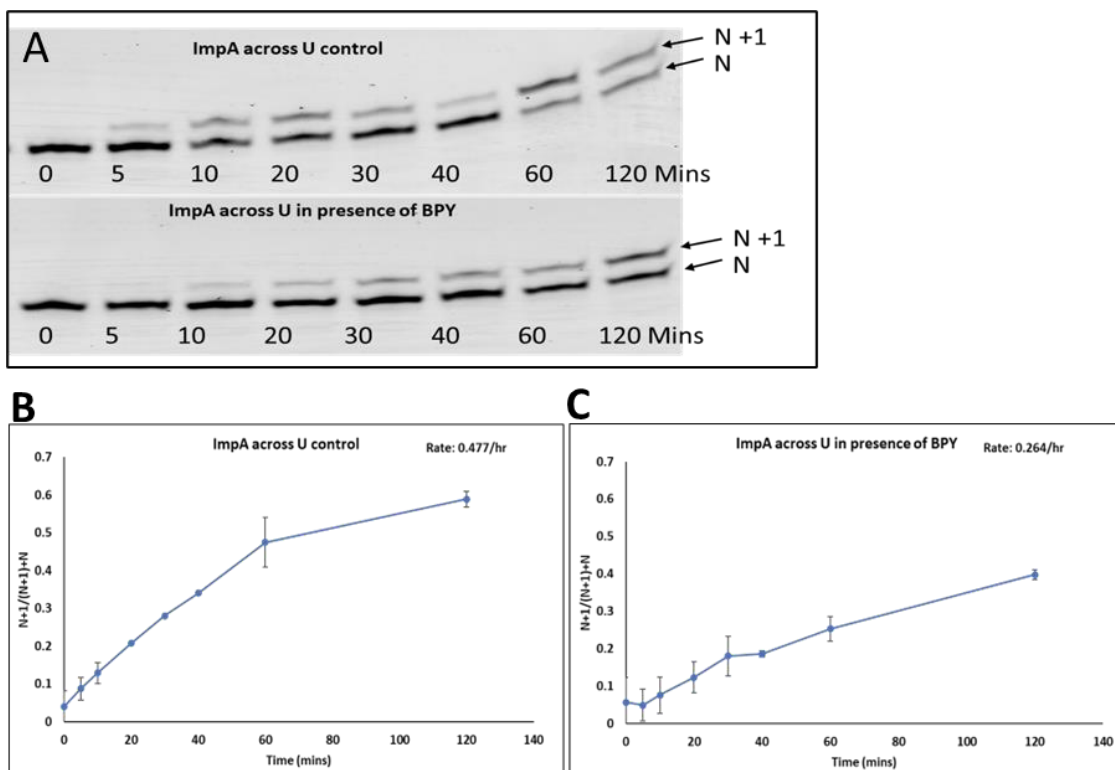


Figure 38: (A) Urea-PAGE representative gel image for control vs test reactions. At specific time points, the samples were taken from the reaction and were analyzed to eventually calculate the extension rate. Rate graph for (B) Control and (C) Test (BPY); the graphs were plotted for the ratio of product formation (on X axis) over time (mins) (Y axis). The rate was calculated by fitting a straight line for the first five time points. (N=2).

### **Studying the formation of complex using UV/Vis analysis**

In order to understand possible interaction happening in the aforementioned reaction between BPY and 5' NMPs, we used UV/Vis spectroscopy to study these complexes as they would show a shift in the pattern of the wavelength when interacting. Upon forming the complex, there is a red shift in the spectrum because of the stabilization of the excited state of the molecules. The red shift was significant in the purines when compared to the pyrimidines. As 5' NMPs and BPY have aromatic systems, they can readily stack on each other forming these complexes, which were shown in previous studies using UV/Vis spectroscopy and potentiometric titration (Griesser et al., n.d.).

A noticeable change in absorption is often observed when two aromatic systems come together to form charge-transfer complexes. UV/Vis spectroscopy is widely used to study this kind of phenomenon and has been reported previously for AMP and BPY. Therefore, we also decided to use UV/VIS-spectroscopy to measure such phenomena in our 5' NMPs and BPY systems. The 5' NMP -BPY system hinted towards forming an adduct when UV/Vis analysis was carried out at the following concentrations, [5' NMP]=10mM [BPY]=0.8mM, in 100mM Tris buffer. The system showed a red shift that can be potentially correlated to adduct formation between AMP/GMP and BPY. The other two nucleotides (UMP and CMP) did not show any significant shift in the spectra. Pi-pi stacking is an attractive, noncovalent interaction (orbital overlap) between the pi bonds of aromatic rings. The interaction occurs within aromatic ring systems, which results in stabilization of the molecule's excited state. This stabilization causes a reduction in the energy difference between the ground and excited states, leading to a longer wavelength and less energy. This phenomenon is known as red shift, which we can observe for the BYP interactions with AMP and GMP in their respective graphs.

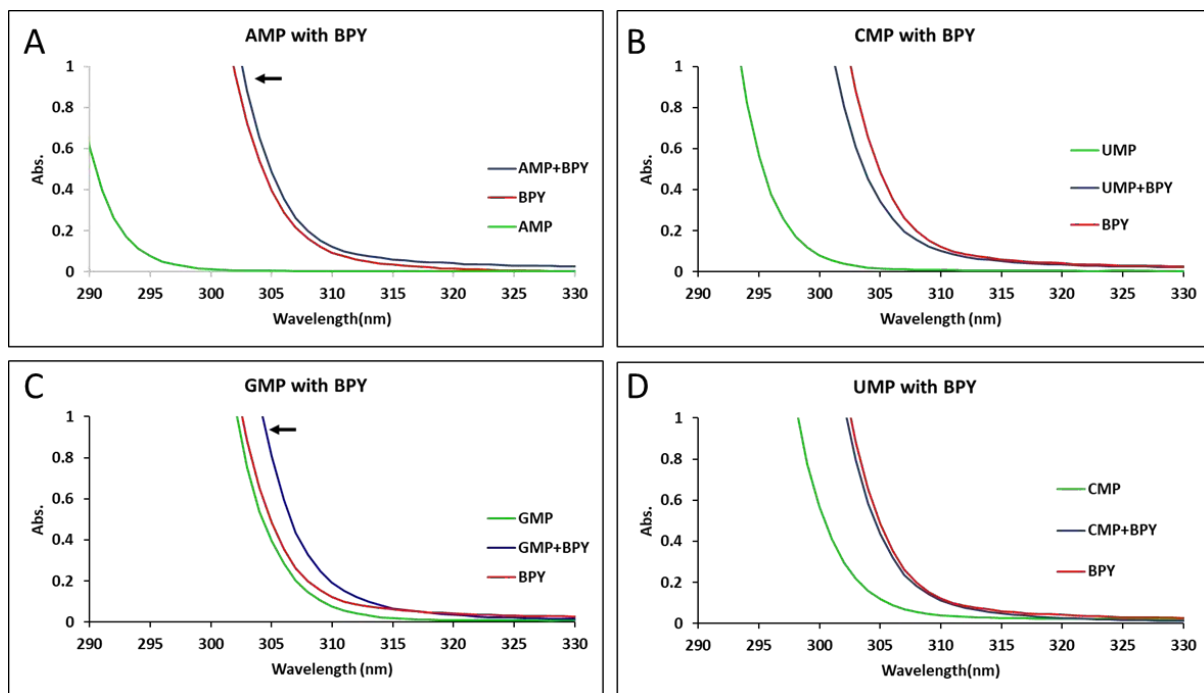


Figure 39: The zoomed in view of 5' NMP -BPY, BPY and 5' NMP spectra taken originally between 200-600nm, Absorbance is plotted on Y axis and wavelength is plotted on X axis. (A) AMP-BPY, BPY, AMP (B) CMP-BPY, BPY, CMP (C) GMP-BPY, BPY, GMP (D) UMP-BPY, BPY, UMP UV/Vis absorbance v/s wavelength at pH=7-7.7 (N=3).

## Chapter 4: Conclusion and Outlook

Due to their intrinsic properties and molecular structure that consists of three different functional groups (nucleobase, sugar and phosphate backbone), nucleotides can self-assemble and interact with surrounding molecules. These dynamic assemblies would have influenced various prebiotic processes thus leading to interesting consequences for the chemistry that panned out on prebiotic Earth. While the self-assembly of nucleotides has been studied, the dynamicity of the assembly has not been that well explored. To address this, we used various spectroscopy tools to understand this phenomenon, especially from a prebiotic context.

To understand the formation of higher-order structures, we first characterized it using Rayleigh scattering. As nucleotide concentration increased, scattering intensity also increased, suggesting the formation of higher order structures. Next, we used pyrene to understand the dynamicity of the system. Peak 1 ( $I_1$ ) and Peak 3 ( $I_3$ ) were used to understand the hydrophobicity, and Peak 1 ( $I_1$ ) and Peak Ex ( $I_{ex}$ ) were used to understand the dynamicity. All the 5`NMPs showed an increase in hydrophobicity. Pyrimidines had relatively more eximerisation than purines, suggesting that their assemblies were more dynamic, whereas the purines showed little to very little excimerization. Further, there was an increase in the anisotropy value with an increase in the concentration. Overall, UMP showed the highest anisotropy value followed by CMP, GMP and AMP, suggesting small size of Ump aggregates which leads to restricted movement of the pyrene molecules.

Upon investigating the temperature dependence of nucleotide self-assembly, a decrease in  $I_{ex}/I_1$  ratio was observed for the pyrimidines (UMP and CMP). This was expected as higher temperatures can lead to more vibrational displacement of molecules, thus destabilizing the structures. In the case of AMP, there was no change in the  $I_1/I_3$  ratio with increasing temperature, indicating highly stable higher-order structures. In the case of GMP, an increase in the  $I_1/I_3$  was observed, which could be due to the breakdown of the higher-order structures, leading to the formation of transient and smaller hydrophobic pockets.

The prebiotic soup would have had a plethora of molecules that could have different effects on the self-assembly of nucleotides, ranging from metal ions to long chain

polymers. To explore this, we studied the effect of prebiotically relevant molecules (metal ions and PEG, a proxy for a prebiotic oligomer), on the self-assembly of nucleotides. Overall, the hydrophobicity was observed to increase in the presence of PEG or metal ions. In the presence of PEG, the hydrophobicity was overall higher to begin with as compared to the without-PEG controls, which got further enhanced upon increasing concentration. In case of metal ions, the hydrophobicity of purines went down and for pyrimidines the hydrophobicity was similar to control. The excimerisation went up in case of purines and went down in case of pyrimidines as we increase the concentration of nucleotides. These results indicated that the presence of oligomeric co-solute molecules (like PEG) and metal ions aid the formation of self-assembled structures of NMPs. In order to understand the composition of these higher-order structures, we used NMR spectroscopy. The H1 NMR showed a shift in the nucleobase protons of the spectra, due to the plausible stacking of the nucleotides, which leads to this up-field shift in the spectra. 2D NOESY NMR confirmed that nucleotides are interacting in specific orientations.

After characterization of the self-assembled structures, we wanted to quickly check the implications of such potential self-assembly to prebiotic chemistry. We used an FDA assay (which is a hydrophobic dye) to investigate how these structures affect the solubilization of hydrophobic molecules. It is important to note that in the prebiotic Earth, the majority of the planet is thought to have been covered with oceans, wherein hydrophobic molecules would have been unable to dissolve in such aqueous conditions. However, we found that the addition of nucleotides led to an increase in the solubility of FDA, as characterized by UV absorbance. In addition, we also investigated how the presence of the aromatic molecule bipyridine (BPY), would have affected nonenzymatic template-directed replication reaction; an important reaction central to emergence of functional RNA molecules. BPY has a tendency to stack with nucleotides thus aiding their self-assembly. Interestingly, we found that the rate of the reaction that we looked at, was significantly reduced when BPY was added to it.

To understand this interaction further, we conducted preliminary UV experiments. The data suggested that the presence of BPY caused a shift in the spectra of purines, indicating that the molecules were involved in stacking-type interactions, while in pyrimidines the shift was not prominent which can be explained by lesser

stacking capabilities. Overall, our results highlighted the intrinsic properties of nucleoside-mono phosphates to self-assemble, the effect of two prebiotically relevant co-solute molecules on this self-assembly, and its implications in two different prebiotically relevant processes (solubilization of hydrophobic molecules and nonenzymatic template-directed replication). We have characterized and confirmed the formation of higher-order structures using UV/Vis, fluorescence spectroscopy, Rayleigh Scattering and Nuclear Magnetic Resonance (NMR). It would be interesting to investigate the effect of other prebiotically relevant co-solute molecules (such as lipids, amino acids etc.), on the self-assembly properties of various NMPs. Pertinently, delineating the structural heterogeneity of these self-assembled structures using sensitive techniques such as X-ray diffraction and electron microscopy, would further shed light on the biochemical underpinnings of this process. This study thus also paves the path for further exploration of self-assembly of nucleotides, and its downstream effects for related biochemical events in extant biological processes.

## References

- Ashbourn, Samantha F M, Jamie E Elsila, Jason P Dworkin, Max P Bernstein, Scott A Sandford, and Louis J Allamandola. 2007. "Ultraviolet Photolysis of Anthracene in H<sub>2</sub>O Interstellar Ice Analogs: Potential Connection to Meteoritic Organics." *Meteoritics & Planetary Science*. Vol. 42. <http://meteoritics.org>.
- Bapat, Niraja V., and Sudha Rajamani. 2015. "Effect of Co-Solutes on Template-Directed Nonenzymatic Replication of Nucleic Acids." *Journal of Molecular Evolution* 81 (3–4): 72–80. <https://doi.org/10.1007/s00239-015-9700-1>.
- Barenholz, Yechezkel, Tina Cohen, Elisha Haas, and Michael Ottolenghi. 1996. "Lateral Organization of Pyrene-Labeled Lipids in Bilayers as Determined from the Deviation from Equilibrium between Pyrene Monomers and Excimers (\*)." *Journal of Biological Chemistry* 271 (6): 3085–90. <https://doi.org/10.1074/JBC.271.6.3085>.
- Burcar, Bradley T., Mohsin Jawed, Hari Shah, and Linda B. McGown. 2015. "In Situ Imidazole Activation of Ribonucleotides for Abiotic RNA Oligomerization Reactions." *Origins of Life and Evolution of Biospheres* 45 (1): 31–40. <https://doi.org/10.1007/s11084-015-9412-y>.
- Butlerow, A. 1861. "Bildung einer zuckerartigen Substanz durch Synthese." *Annalen der Chemie und Pharmacie* 120 (3): 295–98. <https://doi.org/10.1002/jlac.18611200308>.
- Cassidy, Lauren M., Bradley T. Burcar, Wyatt Stevens, Elizabeth M. Moriarty, and Linda B. McGown. 2014. "Guanine-Centric Self-Assembly of Nucleotides in Water: An Important Consideration in Prebiotic Chemistry." *Astrobiology* 14 (10): 876–86. <https://doi.org/10.1089/ast.2014.1155>.

- Dagar, Shikha, Susovan Sarkar, and Sudha Rajamani. 2020. "Geochemical Influences on Nonenzymatic Oligomerization of Prebiotically Relevant Cyclic Nucleotides." *RNA* 26 (6): 756–69. <https://doi.org/10.1261/rna.074302.119>.
- Deamer, David. 2016. "Membranes and the Origin of Life: A Century of Conjecture." *Journal of Molecular Evolution* 83 (5): 159–68. <https://doi.org/10.1007/s00239-016-9770-8>.
- Deamer, David, Sara Singaram, Sudha Rajamani, Vladimir Kompanichenko, and Stephen Guggenheim. 2006. "Self-Assembly Processes in the Prebiotic Environment." *Philosophical Transactions of the Royal Society B: Biological Sciences* 361 (1474): 1809–18. <https://doi.org/10.1098/rstb.2006.1905>.
- Ferris, James P. 1993. "Catalysis and Prebiotic RNA Synthesis." *Origins of Life and Evolution of the Biosphere* 23 (5–6): 307–15. <https://doi.org/10.1007/BF01582081>.
- Ferris, James P., Robert A. Sanchez, and Leslie E. Orgel. 1968. "Studies in Prebiotic Synthesis: III. Synthesis of Pyrimidines from Cyanoacetylene and Cyanate." *Journal of Molecular Biology* 33 (3): 693–704. [https://doi.org/10.1016/0022-2836\(68\)90314-8](https://doi.org/10.1016/0022-2836(68)90314-8).
- Griesser, Rolf, Helmut Sigel, Helmut Sigel<sup>3</sup>, P R Huber, R Griesser, B Prijs, H Sigel, and European J Biochem. n.d. "Ternary Complexes in Solution. VIII. Complex Formation between the Copper(II)-2,2'-Bipyridyl 1:1 Complex and Ligands Containing Oxygen and/or Nitrogen as Donor Atoms<sup>12</sup>." <https://pubs.acs.org/sharingguidelines>.
- Groen, Joost, David W. Deamer, Alexander Kros, and Pascale Ehrenfreund. 2012. "Polycyclic Aromatic Hydrocarbons as Plausible Prebiotic Membrane Components." *Origins of Life and Evolution of Biospheres* 42 (4): 295–306. <https://doi.org/10.1007/s11084-012-9292-3>.
- Hud, Nicholas V., Brian J. Cafferty, Ramanarayanan Krishnamurthy, and Loren Dean Williams. 2013. "The Origin of RNA and 'My Grandfather's Axe.'" *Chemistry and Biology* 20 (4): 466–74. <https://doi.org/10.1016/j.chembiol.2013.03.012>.
- Izgu, Enver Cagri, Albert C. Fahrenbach, Na Zhang, Li Li, Wen Zhang, Aaron T. Larsen, J. Craig Blain, and Jack W. Szostak. 2015. "Uncovering the Thermodynamics of Monomer Binding for RNA Replication." *Journal of the American Chemical Society* 137 (19): 6373–82. <https://doi.org/10.1021/jacs.5b02707>.
- Kaddour, Hussein, Selim Gerislioglu, Punam Dalai, Toshikazu Miyoshi, Chrys Wesdemiotis, and Nita Sahai. 2018. "Nonenzymatic RNA Oligomerization at the Mineral–Water Interface: An Insight into the Adsorption–Polymerization Relationship." *The Journal of Physical Chemistry C* 122 (51): 29386–97. <https://doi.org/10.1021/acs.jpcc.8b10288>.
- Kaddour, Hussein, and Nita Sahai. 2014. "Synergism and Mutualism in Non-Enzymatic RNA Polymerization." *Life* 4 (4): 598–620. <https://doi.org/10.3390/life4040598>.
- Liang, Hao, Feifei Lin, Zijie Zhang, Biwu Liu, Shuhui Jiang, Qipeng Yuan, and Juewen Liu. 2017. "Multicopper Laccase Mimicking Nanozymes with Nucleotides as Ligands." *ACS Applied Materials and Interfaces* 9 (2): 1352–60. [https://doi.org/10.1021/ACSAMI.6B15124/ASSET/IMAGES/LARGE/AM-2016-151242\\_0007.JPEG](https://doi.org/10.1021/ACSAMI.6B15124/ASSET/IMAGES/LARGE/AM-2016-151242_0007.JPEG).
- Lopez, Anand, and Juewen Liu. 2017. "Self-Assembly of Nucleobase, Nucleoside and Nucleotide Coordination Polymers: From Synthesis to Applications." *ChemNanoMat* 3 (10): 670–84. <https://doi.org/10.1002/cnma.201700154>.
- Mahajan, Tania B., Jamie E. Elsil, David W. Deamer, and Richard N. Zare. 2003. "Formation of Carbon-Carbon Bonds in the Photochemical Alkylation of Polycyclic Aromatic Hydrocarbons." *Origins of Life and Evolution of the Biosphere* 33 (1): 17–35. <https://doi.org/10.1023/A:1023996314942>.



- Miller, Stanley L. 1953. "A Production of Amino Acids Under Possible Primitive Earth Conditions." *Science* 117 (3046): 528–29. <https://doi.org/10.1126/science.117.3046.528>.
- Morowitz, Harold J., Bettina Heinz, and David W. Deamer. 1988. "The Chemical Logic of a Minimum Protocell." *Origins of Life and Evolution of the Biosphere* 18 (3): 281–87. <https://doi.org/10.1007/BF01804674>.
- Mungi, Chaitanya V., Niraja V. Bapat, Yayoi Hongo, and Sudha Rajamani. 2019. "Formation of Abasic Oligomers in Nonenzymatic Polymerization of Canonical Nucleotides." *Life* 9 (3). <https://doi.org/10.3390/life9030057>.
- Nelson, Kevin E., Matthew Levy, and Stanley L. Miller. 2000. "Peptide Nucleic Acids Rather than RNA May Have Been the First Genetic Molecule." *Proceedings of the National Academy of Sciences* 97 (8): 3868–71. <https://doi.org/10.1073/pnas.97.8.3868>.
- Oba, Yasuhiro, Yoshinori Takano, Yoshihiro Furukawa, Toshiki Koga, Daniel P. Glavin, Jason P. Dworkin, and Hiroshi Naraoka. 2022. "Identifying the Wide Diversity of Extraterrestrial Purine and Pyrimidine Nucleobases in Carbonaceous Meteorites." *Nature Communications* 13 (1): 2008. <https://doi.org/10.1038/s41467-022-29612-x>.
- Odom, D. G., M. Rao, J. G. Lawless, and J. Oro. 1979. "Association of Nucleotides with Homoionic Clays." *Journal of Molecular Evolution* 12 (4): 365–67. <https://doi.org/10.1007/BF01732031>.
- Orgel, Leslie. 2000. "A Simpler Nucleic Acid." *Science* 290 (5495): 1306–7. <https://doi.org/10.1126/science.290.5495.1306>.
- Oro, J. 1995. "Chemical Synthesis of Lipids and the Origin of Life." *Journal of Biological Physics* 20 (1): 135–47. <https://doi.org/10.1007/BF00700430>.
- Oró, J., and S. S. Kamat. 1961. "Amino-Acid Synthesis from Hydrogen Cyanide under Possible Primitive Earth Conditions." *Nature* 190 (4774): 442–43. <https://doi.org/10.1038/190442a0>.
- Robertson, Michael P., and Gerald F. Joyce. 2012. "The Origins of the RNA World." *Cold Spring Harbor Perspectives in Biology* 4 (5): 1. <https://doi.org/10.1101/cshperspect.a003608>.
- Schlesinger, Gordon, and Stanley L. Miller. 1983. "Prebiotic Synthesis in Atmospheres Containing CH<sub>4</sub>, CO, and CO<sub>2</sub>." *Journal of Molecular Evolution* 19 (5): 376–82. <https://doi.org/10.1007/BF02101642>.
- Sigel, Astrid, Bert P. Operschall, and Helmut Sigel. 2014. "Comparison of the  $\pi$ -Stacking Properties of Purine versus Pyrimidine Residues. Some Generalizations Regarding Selectivity." *Journal of Biological Inorganic Chemistry* 19 (4–5): 691–703. <https://doi.org/10.1007/s00775-013-1082-5>.
- Sigel, Helmut. 2004. "Adenosine 5'-Triphosphate (ATP 4-): Aspects of the Coordination Chemistry of a Multitalented Biological Substrate\*." *Pure Appl. Chem.* Vol. 76.
- Stokkeland, Ingrid, and Peter Stilbs. 1985. "A Multicomponent Self-Diffusion NMR Study of Aggregation of Nucleotides, Nucleosides, Nucleic Acid Bases and Some Derivatives in Aqueous Solution with Divalent Metal Ions Added." *Biophysical Chemistry* 22 (1–2): 65–75. [https://doi.org/10.1016/0301-4622\(85\)80026-0](https://doi.org/10.1016/0301-4622(85)80026-0).
- Szostak, Jack W. 2012. "The Eightfold Path to Non-Enzymatic RNA Replication." *Journal of Systems Chemistry* 3 (1). <https://doi.org/10.1186/1759-2208-3-2>.
- Wong, Alan, Ramsey Ida, Lea Spindler, and Gang Wu. 2005. "Disodium Guanosine 5'-Monophosphate Self-Associates into Nanoscale Cylinders at PH 8: A Combined Diffusion NMR Spectroscopy and Dynamic Light Scattering Study." *Journal of the American Chemical Society* 127 (19): 6990–98. <https://doi.org/10.1021/ja042794d>.
- Xu, Huifang, Na Du, Yawen Song, Shue Song, and Wanguo Hou. 2018. "Spontaneous Vesicle Formation and Vesicle-to-Micelle Transition of Sodium 2-Ketoacetate in

- Water.” *Journal of Colloid and Interface Science* 509 (January): 265–74.  
<https://doi.org/10.1016/J.JCIS.2017.09.023>.
- Yadav, Mahipal, Ravi Kumar, and Ramanarayanan Krishnamurthy. 2020. “Chemistry of Abiotic Nucleotide Synthesis.” *Chemical Reviews* 120 (11): 4766–4805.  
<https://doi.org/10.1021/acs.chemrev.9b00546>.
- Zhou, Huan Xiang, Germán Rivas, and Allen P. Minton. 2008. “Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences.” *Annual Review of Biophysics* 37: 375–97.  
<https://doi.org/10.1146/annurev.biophys.37.032807.125817>.
- Zhou, Li, Zhenhua Li, Zhen Liu, Meili Yin, Jinsong Ren, and Xiaogang Qu. 2014. “One-Step Nucleotide-Programmed Growth of Porous Upconversion Nanoparticles: Application to Cell Labeling and Drug Delivery.” *Nanoscale* 6 (3): 1445–52.  
<https://doi.org/10.1039/C3NR04255C>.