Effect of various co-solutes on the oligomerization of prebiotically relevant monomers

Indian Institute of Science Education and Research, Pune

Thesis submitted in partial fulfillment of the requirements of Five year BS-MS Dual Degree Program

By

Aparna P.K.

20111037

Biology

Under the guidance of

Dr. Sudha Rajamani

Asst. Professor (Biology)

IISER Pune

Certificate

This is to certify that this dissertation entitled "**Effect of various co-solutes on the oligomerization of prebiotically relevant monomers**", towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune, represents the research carried out by "**Aparna P K at IISER Pune**" under the supervision of "**Dr Sudha Rajamani**, **Asst. Professor, Biology**," during the academic year 2015-2016.

Dr. Sudha Rajamani 28-03-16

Declaration

I hereby declare that the matter embodied in the report entitled “**Effect of various co-solutes on the oligomerization of prebiotically relevant monomers**” are the results of the investigations carried out by me at the Department of Biology, IISER Pune, under the supervision of Dr. Sudha Rajamani and the same has not been submitted elsewhere for any other degree.

Aparna P.K. 28-03-16

20111037

Abstract

RNA world hypothesis suggests that RNA was the first informational biopolymer from which DNA and proteins evolved later. However, this hypothesis has historically ignored the role of other prebiotically relevant molecules that would have played an important role in the shaping of the RNA world. According to this hypothesis, synthesis of short RNA oligomers, from a large random pool of organic molecules, is thought to have happened by chemical means. Dehydration- rehydrations (DH-RH), regimes, a common theme on the prebiotic Earth, are thought to have especially favoured such uphill oligomerization reactions on the early Earth. DH-RH cycling of ribonucleotides has shown to result in RNA-like oligomers in the presence of lipids. In the current project, we set out to characterize the combined role of prebiotically relevant co-solutes, such as lipids and amino acids (for e.g. glycine, alanine, aspartate and valine), on the oligomerization of non-activated nucleotides under DH-RH cycling conditions. The aforementioned moieties are important, prebiotically relevant co-solutes that would have co-existed with the RNA monomers in the prebiotic soup. We also characterized the role of these co-solutes on depurination; an important side-reaction that has implications for the origin of informational polymers on the early Earth. In addition to this, we also wanted to understand the role of these co-solutes on the oligomerization reactions of activated nucleotides as most studies of nonenzymatic oligomerization involved the use of these activated monomers. In all, our results suggest that there is no additive role resulting from combining lipids and amino acids on, both, the oligomerization and depurination of non-activated nucleotides. Activated nucleotides also did not oligomerize efficiently in the presence of the lipids and/or amino acids under our reaction conditions. Importantly, the presence of lipids did confer some protection from depurination towards nucleotides in all the reactions that we studied.

Table of Contents

Certificate 1

Declaration 2

Abstract 3

List of Tables: 6

List of figures: 7

Acknowledgement 8

Background literature: 9

Questions we addressed: 11

Materials: 14

Methods: 14

Preparation of stocks solutions: 14

Preparation of lipid stock solutions: 14

Preparation of lipid vesicles: 15

Reaction conditions: 15

1. DH-RH Cycling reaction for non-activated nucleotides: 15
2. Reactions involving activated nucleotides: 16

Extraction of lipid from reaction time points: 18

High Performance liquid chromatography (HPLC): 18

Matrix-assisted laser desorption/ionization spectrometry (MALDI): 19

Matrix preparation: 20

Sample preparation: 20

High resolution mass spectrometry (HRMS): 20
Results and discussion: 21
1a. Oligomerization of AMP in the presence of DLPC and amino acids 21
1b. Depurination in DH-RH cycling reactions involving non-activated nucleotides: 24
2. Activated nucleotide oligomerization under different prebiotic conditions: 30
a) Role of DLPC and alanine as co-solutes on the oligomerization of ImpG at room temperature. 30
b) Dehydration reaction of ImpG in the presence of DLPC and alanine: 32
c) Oligomerization of ImpG in the presence of DMPC/POPC under vacuum at room temperature: 32
d) Oligomerization of ImpG in the presence of DMPC at room temperature: 33
References: 34

List of Tables:

Table 1: List of DH-RH reactions using DLPC and various amino acids 15
Table 2: ImpG reactions with Ala and DLPC as cosolutes. 16
Table 3: ImpG centrivap experiment in the presence of DLPC and Ala 17
Table 4: HPLC gradients for (a) non activated nucleotide (b) activated nucleotide at 0.8ml/min flow rate. 19
Table 5: Mass obtained from MALDI from Control and reaction samples. 27
Table 6: Standardization of control samples for non activated and activated nucleotide reactions using HRMS. 30

List of figures:

Figure 1: HPLC chromatogram of AMP oligomerization in the presence of DLPC and Asp over cycles 0, 1, 4 and 7, ... 21
Figure 2: HPLC chromatogram comparison of main reaction (AMP with DLPC and Asp) along with the control reactions. 22
Figure 3: HPLC chromatogram comparison of AMP oligomerization in the presence of Gly, Ala, Asp and Val. 23
Figure 4: HPLC chromatogram comparison of AMP oligomerization in the presence of both lipid plus various amino acids. 24
Figure 5: Representative HPLC chromatogram showing depurination during AMP oligomerization in the presence of DLPC and amino acids. 25
Figure 6: Quantitative area analysis of HPLC chromatograms of the control reactions and Asp reaction (cycle 4). 25
Figure 7: Comparison of quantitative area analysis of a) only amino acids b) amino acids+ lipid c) Zoomed view of oligomers in 7b. 26
Figure 10: HPLC chromatograms of ImpG, ImpG+ DLPC, ImpG+ Ala and ImpG+ DLPC+ Ala 5th day time point. 31
Figure 11: HRMS standardization of a) Poly G_4 b) Poly G_5 31
Figure 12: Comparison of ImpG, ImpG+ DMPC and ImpG+ POPC 5th day time point with ImpG and GMP control. 32
Figure 13: HPLC chromatogram of 2nd day time point of ImpG and ImpG+ DMPC. 33

Acknowledgement

I would like to express my sincere gratitude towards Dr Sudha Rajamani for giving me this opportunity to work with her on my Master's thesis. I am grateful to her for her guidance, encouragement and continuous support throughout the project.

I am greatly thankful to Chaitanya for his countless helps and valuable suggestions for my project. I would like to thank Niraja for doing MALDI for all my samples and giving ideas and suggestions. I am also thankful to Rajeswari for timeless helps. My thanks also go to other lab mates Pratima, Susovan and Rutwik.

I would like to thank Dr. Ajeet Singh and CAMS Venture center for carrying out HRMS characterization of the samples.

I would like to thank IISER Pune for giving me an opportunity to study this research oriented course and INSPIRE scholarship for providing funding support.

Last but not the least I am grateful to my parents and my sisters for their love, faith and support throughout these many years.

Background literature:

The sequence of events that led to the origin of life on Earth is still an unanswered question. Abiogenesis is one of the important processes by which life may have originated on the early Earth and the Urey-Miller experiment provided the footstone for this theory (Miller, 1953). According to this view, biologically relevant monomers and oligomers chemically originated, and the oligomers subsequently evolved due to their self-replicating capabilities. It is thought that they used chemical means to pass on their information to subsequent generations and assembled into complex entities, along with membranous molecules (e.g. fatty acids), leading to the origin of primitive cells. It is still debated as to which set of molecules might have evolved first. Several ideas have been proposed and studied regarding the chemical origins of life. Thus far, the RNA World hypothesis (Gilbert, 1986) happens to be the most accepted of them all. This hypothesis proposes that RNA could have been the first informational biomolecule to emerge from relevant prebiotic chemistry that also catalyzed chemical reactions. Subsequently, DNA and proteins are thought to have evolved from RNA by a process of chemical evolution that relegated each to their specific roles they play in extant biochemistry.

Chemical evolution towards an RNA world can be divided into three processes: formation of the precursors of the RNA i.e. ribonucleotides, oligomerization of these monomers to result in oligomers, and, finally, the evolution of these to result in catalytically active ribozymes. The formation of nucleotides under prebiotically relevant conditions has been demonstrated by a few groups. Despite the fact that the complexity of these molecules and their thermodynamic instability in aqueous solution make their synthesis a non-trivial task, recent studies from Sutherland and co-workers have demonstrated the possibility of the formation of various biomolecular precursors relatively easily, including that of nucleotides, under prebiotically relevant early Earth conditions. Synthesis of pyrimidine and purine nucleotide precursors has been previously reported (Powner et al., 2009; Powner et al., 2010). Formation of ribose, ribonucleosides and short RNAs from atmospheric CO_2, H_2O and N_2 has also been demonstrated experimentally (Benner et al., 2012). More recent studies have shown that reductive homo ligation of hydrogen cyanide and its derivatives leads to the synthesis of ribonucleotides precursors (Patel et al., 2015). In one study, high energy

proton irradiation of liquid formamide, in the presence of meteorite, was shown to result in the formation different prebiotically relevant molecules including different nucleobases and nucleosides (Saladino et al., 2015). All the aforementioned studies indicate the various routes by which ribonucleotides and related molecules could have been synthesized on the early Earth.

One of the milestone discoveries in the history of the RNA World hypothesis was made by the 1982 Nobel Prize winning study that demonstrated the catalytic ability of RNA molecules (Stark et al., 1978; Zaug and Cech, 1980). It is thought that a minimal length of 30-50 nucleotides of RNA is required for the molecule to be able to become a functional ribozyme. The nonenzymatic evolution of a fully functioning ribozyme of this length, from a random population of various chemical molecules, must have been a difficult process. Therefore, there should have been some chemical means which would have made the phosphodiester bond formation thermodynamically and kinetically feasible. Towards this end, several means of oligomerization of ribonucleotides have been characterized. In 1951, Bernal and Goldschmidt proposed that clay minerals could have played an important role in the oligomerization of prebiotically relevant monomers. In the presence of montmorillonite clay, imidazole-activated monomers of RNA formed oligomers up to 6-14 nucleotides long in aqueous solution (Ferris and Ertem, 1993). Activation with 1-methyl adenine resulted in the formation RNA oligomers of 40-50 nucleotides in length (Huang and Ferris, 2003). Additionally, eutectic phases in ice have also been considered as one of the plausible scenarios where prebiotic chemistry could have occurred. Eutectic temperature of water is -18.4^{o}C, which has been shown to favour the oligomerization reaction of activated nucleotides (Kanavarioti et al., 2001). Importantly, in all the aforementioned studies, activation chemistry had been used to facilitate condensation reactions of monomers in aqueous conditions. However, the prebiotic synthesis of activated ribonucleotides has not been convincingly demonstrated yet. This makes acidic geothermal pools a relevant candidate for studying the oligomerization of non-activated nucleotides as the inherent chemical and thermal fluxes present in these scenarios have been shown to facilitate uphill condensation reactions. The temperature fluctuation due to day-night cycles leads to continuous dehydration-rehydration (DH-RH) of the pools. High temperatures help the reactants to

overcome thermodynamic barrier; there is reduced water activity which also enables concentrating of the monomers, all of which facilitate the formation of oligomers. Acidic pH can facilitate acid catalyzed phosphodiester bond formation similar to the acid-mediated ester bond formation. Oligomerization of different molecules such as alpha hydroxy acids, malic acid etc under wet dry cycles had been reported (Mamajanov et al., 2014; Weber, 1989). Importantly, when nucleoside-5'-monophosphates were subjected to DH-RH cycles in the presence of phospholipids, RNA-like oligomers up to 50-75 nucleotides were shown to be formed (Rajamani et al., 2008).

Questions we addressed:

All of the aforementioned reactions discussed thus far have overlooked the fact that the prebiotic soup would have been a heterogeneous mix of chemical moieties as against a pure system consisting of only one kind of special molecule. Only one study to date has characterized the role of co-solutes on nonenzymatic template-directed replication of nucleic acids (Bapat and Rajamani, 2015). This study clearly demonstrated the importance of considering the molecular complexity of the prebiotic soup while studying prebiotically relevant reactions. Therefore, we thought it pertinent to include other prebiotically relevant molecules in our oligomerization reactions to get a realistic sense of the role that the presence of these co-solutes might have had on our reactions. In this context, we studied the effect of lipids and amino acids, as they are very important co-solute molecules, which also play crucial roles in contemporary biology.

Synthesis and abundance of complex phospholipids on the prebiotic Earth is debatable. Instead, it is thought that fatty acids were components of prebiotic amphiphilic compartments. However, given that phospholipids are the major contemporary amphiphiles and have also been shown to form stable vesicles over wide pH and temperature regimes, we used them in this preliminary study to gain a better understanding of the role of amphiphiles as co-solutes in early chemical reactions. Amino acids, on the other hand, are prebiotically relevant molecules and the synthesis of most of the naturally occurring amino acids has been demonstrated under prebiotic conditions (Miller, 1953; Patel et al., 2015). Very few studies to date have looked at their

role in the shaping of the RNA World. In one relevant study, histidine containing di-peptides were shown to facilitate phosphodiester bond formation by covalently linking to the activated nucleotide and forming an intermediate in the process (Wieczorek et al., 2013). In this study, we set out to study the role of lipids and different amino acids on the oligomerization of, both, activated and non-activated nucleotides, under prebiotically relevant conditions.

*Oligomerization of non-activated nucleotides:*1, 2-dilauroyl-sn-glycero-3-phosphocholine (DLPC), glycine (Gly), alanine (Ala), aspartate (Asp) and valine (Val) - (GADV amino acids), were used in this context to understand their role on the oligomerization of non-activated nucleotides under DH-RH cycling conditions. On dehydration, lipids are known to form multilamellar phases that can concentrate reactants and favour oligomerization. On the other hand, under DH-RH conditions, amino acids might also be able to form covalent bonds with nucleotides, which, we hypothesized, might facilitate phosphodiester bond formation as was shown in the aforementioned study. The amino acids GADV were chosen randomnly. Additionally, under volcanic geothermal conditions, nonenzymatic depurination of the nucleotides is commonly observed (An et al., 2014). This leads to the loss of information of the oligomer due to losing its nitrogenous base(s). However, it has been shown that phospholipids can not only act as concentrating agents, but can also provide a protecting environment to nucleotides from depurination (Mungi and Rajamani, 2015). In this study, we also wanted to see whether lipid and amino acids might have any additive/synergetic role on protection of the information of the nucleotides under low pH and high temperature conditions.

Oligomerization of activated nucleotides: We also studied the role of presence of alanine, DLPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and/or 1, 2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) on the oligomerization of imidazole-activated nucleotides at room temperature. We chose alanine as the amino acid for this study as it facilitated relatively greater oligomerization in the study involving non-activated nucloetides. While the role of lipids was alluded to earlier, to date, no study has been undertaken to study the potential catalytic role of lipids on the oligomerization

of activated nucleotides, which was our original aim. Furthermore, amino acids are thought to be interacting via their amino (NH_3^+) group with the phosphate end of the nucleotide, thus possibly stabilizing the nucleotide. Therefore, we also studied the combined role of lipid and amino acids on the oligomerization of activated nucleotides in this context. We undertook these sets of reactions at both moderate and subzero temperatures (to study eutectic phase reaction).

Additionally, recent X-ray diffraction data from Dr Raghunathan's lab at the Raman Research Institute (Bangalore) showed a very interesting observation for a concentrated mixture of adenosine monophosphate (AMP, 10mg/mL) and the lipid DMPC. As the transition temperature of DMPC is 24°C, the lipid is in a gel-liquid form at 24°C. But, when AMP is present in either 1:1, 2:1 or 4:1 ratio to the lipid, an aqueous solution with two phases was observed between 24-50°C (personal communication). One of these phases is thought to be enriched in mononucleotides. It is hypothesized that the nitrogenous base gets inserted into the lipid bilayer in a manner similar to cholesterol insertion into the membranes. This should potentially leave the phosphate backbone in the water interlayer in close proximity to each other, which might facilitate oligomerization. This kind of concentration of nucleotides, we hypothesized, might have implications for the prebiotic oligomerization of imidazole-activated ribonucleotides. Given this, we also carried out experiments using activated nucleotides and DMPC at, both, 24°C and 55°C.

Materials:

Sodium salt of Adenosine mono phosphate (AMP), guanosine mono phosphate disodium salt hydrate (GMP), glycine (Gly), L-alanine (Ala), L-aspartate (Asp), L-valine (Val), acetonitrile, Trizma base, sodium chloride (NaCl), hydro chloric acid (HCl), sulfuric acid (H_2SO_4), butanol, hexane, sodium perchlorate ($NaClO_4$) were purchased from Sigma Aldrich. 1, 2-dilauroyl-sn-glycero-3-phosphocholine (DLPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) were purchased from Avanti polar lipids. Guanosine mono phosphate imidazolide (ImpG) was purchased from GLSynthesis Inc. 5'-AA dimer and 5'-AAA trimer were purchased from Dharmacon part of GE Health care. Poly G_4, Poly G_5 and Poly G_{21} were purchased from Thermo scientific.

Methods:

Preparation of stocks solutions:

1. 100mM AMP stock solution was prepared by dissolving 37 mg in 1 ml milliQ and stored at -20°C.
2. 100mM ImpG stock solution was freshly prepared each time as per the requirement.
3. 25mM stock solutions of 10 ml of all four amino acids was prepared by dissolving 18.8 mg for Gly, 22.3 mg for Ala, 33.3 mg for Asp and 29.3 mg for Val, respectively in milliQ.
4. 2mL of 100mM H_2SO_4 was prepared by adding 27.1 µl of concentrated H_2SO_4.
5. 1M Tris buffer was prepared as stock solution by dissolving 1.21 g of Trizma in 10 ml of milliQ.

Preparation of lipid stock solutions:

In order to prepare 20 mM lipid stock solutions, 12.437 mg, 15.20 mg and 13.55 mg of DLPC, POPC and DMPC, respectively, were taken in glass vials from their 25mg/ml chloroform stocks. After complete evaporation of chloroform, 1 ml milliQ was added to prepare the final solution.

Preparation of lipid vesicles:

In order to start with homogeneous lipid vesicles in reactions involving DLPC and POPC, the final solutions were extruded through a membrane with pore size 800 nm using an extruder from Avanti Polar Lipids Inc. In reactions involving DMPC, the lipid solution was sonicated to simulate the details as in Dr. Raghunatha's experiments.

Reaction conditions:

1. **DH-RH Cycling reaction for non-activated nucleotides:**

Reaction conditions were standardized previously in the lab (Mungi and Rajamani, 2015). Final concentrations of various components in the reactions, depending on the reaction, were as follows: 5mM AMP, 1mM DLPC and 5mM amino acid (AMP: Amino acid in 1:1 mole ratio). For each of the amino acid, four reactions were carried out; AMP alone, AMP+DLPC, AMP+amino acid (all control reactions) and AMP+DLPC+amino acid (the main reaction). Reaction mixtures were rehydrated with milliQ. Each DH-RH cycle was of 1 hr and time points of 200 µl were taken after 0, 1, 4 and 7 cycles, respectively. The details of the reactions and the ingredients are listed in Table 1.

Sl/No	Reaction components	Temperature	pH	Reaction Volume(µl)	Time per cycle	No of cycles
1	AMP	90°C	2	200	1 hr	7
2	AMP+DLPC	90°C	2	200	1 hr	7
3	AMP+ Gly	90°C	2	200	1 hr	7
4	AMP+ Ala	90°C	2	200	1 hr	7
5	AMP+ Asp	90°C	2	200	1 hr	7
6	AMP+ Val	90°C	2	200	1 hr	7
7	AMP+DLPC+Gly	90°C	2	200	1 hr	7
8	AMP+DLPC+Ala	90°C	2	200	1 hr	7
9	AMP+DLPC+Asp	90°C	2	200	1 hr	7
10	AMP+DLPC+Val	90°C	2	200	1 hr	7

Table 1: List of DH-RH reactions using DLPC and various amino acids

2. **Reactions involving activated nucleotides:**

a) **Role of DLPC and alanine as co-solutes on the oligomerization of ImpG at room temperature.**

Final concentrations in the reaction were as follows: 5mM ImpG, 5mM Ala and 1 mM DLPC. pH of the reaction mixture was adjusted to 8 using 5 mM Tris buffer. Reaction volume was kept at 50 µl and time points were taken at 0 min, 12hrs, 1 day, 2 days and 5 days respectively. Samples were analyzed after butanol-hexane extraction using High Performance liquid chromatography (HPLC). List of reactions are given in Table 2.

Sl/No	Reaction components	Temperature	pH	Reaction Volume(µl)
1	ImpG	24°C	8	50
2	ImpG+ DLPC	24°C	8	50
3	ImpG+ Ala	24°C	8	50
4	ImpG+ DLPC+ Ala	24°C	8	50

Table 2: ImpG reactions with Ala and DLPC as cosolutes.

b) **Dehydration reaction of ImpG in the presence of DLPC and Alanine:**

To study the effect of DLPC on the oligomerization of ImpG under simulated DH-RH conditions, reactions were carried out using a speed vac. We wanted to test if dehydrated lipid matrices might also facilitate greater oligomerization of activated nucleotides as there has been no such study reported till date. Since the speed vac could not be controlled for temperature, these reactions occurred at 30°C due to heating up of the instrument's motor. Given the sensitivity of the imidazole group, we were concerned about the stability of ImpG molecule but went ahead with the experiment nevertheless. In one variation of this reaction, Ala was also added. Final concentrations of components in the reactions were 5mM ImpG, 5mM Ala , 1 mM DLPC and 5 mM Tris buffer (pH=8) and reaction volume was 50µl. Each DH-RH cycle was of 45 minutes and time points were taken after 0, 1, 3 and 5 cycles, respectively. Reaction samples were rehydrated using 5mM Tris buffer. Reaction details are listed in Table 3.

Sl/No	Reaction components	Temperature	pH	Reaction Volume(µl)	Time per cycle	No of cycles
1	ImpG	30^0C	8	50	45 min	5
2	ImpG+ DLPC	30^0C	8	50	45 min	5
3	ImpG+ Ala	30^0C	8	50	45 min	5
4	ImpG+ DLPC+ Ala	30^0C	8	50	45 min	5

Table 3: ImpG centrivap experiment in the presence of DLPC and Ala

c) Oligomerization of ImpG In the presence of DMPC at room temperature:

As alluded to in the introduction, oligomerization of ImpG was carried out in the presence of DMPC at room temperature. The ratio of ImpG to DMPC was kept as 2:1 and final concentrations of the reaction mixture were 25mM ImpG and 12.5mM DMPC to simulate reaction conditions similar to those of Dr. Raghunathan's experiment. The reaction was carried out at 24°C at pH 8 (Tris buffer). As a negative control, reaction with GMP was also carried out as there would be no inherent activation energy source for GMP oligomerization in the absence of the imidazole group.

d) Oligomerization of ImpG in the presence of DMPC/POPC under vacuum at room temperature

Reactions were also carried out at 55°C using final concentrations of 2mM ImpG and 1mM DMPC/POPC and pH was adjusted to 8 using 5mM Tris buffer. In these reactions, the DMPC solution was preheated to 55°C and then sonicated to get unilamellar vesicles. POPC was used as a negative control as its transition temperature is -2°C and there should be no phase-separation in this case (personal communication with Dr. Raghunath). Reactions were kept under vacuum to enhance dehydration and, thus, to

promote the condensation reaction. Reaction volumes used were of 200µl and time points were taken after 0hr, 12hrs, 1day, 2days and 5days respectively.

Extraction of lipid from reaction time points:

Butanol-hexane extraction was carried out to remove the lipid from the time point samples that contained them to facilitate biochemical analysis of the aqueous phases. Sample was treated with butanol with 2:1 volume to remove lipid followed by hexane 1:1 volume to remove excess Butanol. Using centri-vap excess hexane was evaporated and volume of each sample was made up to 200µl. This is done because, during lipid extraction, some amount of water can be pulled out by butanol along with the lipid, which results in differential loss of water from the samples. In all cases, the aqueous samples were subjected to further biochemical analysis to check for the presence of oligomeric products.

High Performance liquid chromatography (HPLC):

HPLC is an analytical tool used to separate compounds based on their differential affinity towards the stationary and mobile phases. The column is tightly packed with matrix to form a stationary phase and the mobile phase is passed through the column under high pressure. Samples are eluted at different times because of their difference in relative affinities for the mobile phase and the stationary phase.

In our case, we had used ion exchange HPLC chromatography. This HPLC helps to separate out the molecules depending on the number of charges on it. Anion exchange chromatography was carried out on Agilent 1260 Infinity series (Agilent Technologies, Santa Clara, CA, USA) using a DNAPac PA200 column (4 mm × 250 mm) from Dionex (Thermo Scientific, Sunnyvale, CA, USA).The stationary phase in this case is positively charged so that it can bind to negatively charged molecules. The mobile phase is a mixture of solvent A (2mM Tris pH 8) and solvent B (2mM Tris + 0.4M $NaClO_4$). As the salt is pushed through the column, molecules with least number of negative charges start eluting as their affinity decreases towards the stationary phase. Further increase in

the salt concentration results in the elution of the other components in the order of their increasing negative charge.

Two different salt gradients were used for analysis of resultant products from reactions involving activated and non-activated nucleotides as shown in Table 4.

a)

Time	A%	B%
0	100	0
3.4	100	0
15.4	50	50
16.8	0	100
22.8	0	100
25.2	100	0
31.2	100	0

b)

Time	A%	B%
0	100	0
3.6	100	0
15.8	93	7
20.6	87	13
23	75	25
25.4	60	40
27.8	0	100
35	0	100
37.4	100	0
42.4	100	0

Table 4: HPLC gradients for (a) non activated nucleotide (b) activated nucleotide at 0.8ml/min flow rate.

Each time point from every reaction was analyzed using HPLC. In order to study depurination in a quantitative manner, HPLC runs were also performed with smaller loads to get sharp clean peaks and the graphs were plotted using area% of the relevant peaks from the respective HPLC chromatograms.

Matrix-assisted laser desorption/ionization spectrometry (MALDI):

MALDI is an ionization technique used to mass characterize different biomolecules and large organic molecules. Using laser, molecules are ionized via energy transfer from matrix. Charged ions of various sizes are generated and their time of flight is analyzed (TOF analyzer) based on their mass to charge ratio and corresponding mass is obtained.

Matrix preparation:

500 mM 2′, 4′, 6′-Trihydroxyacetophenone (THAP) is dissolved in Ethanol. To the THAP solution 0.1M solution of Ammonium hydrogen citrate is added in 2:1 ratio.

Sample preparation:

Reaction samples had been collected as separated peaks using anion exchange chromatography. After lyophilization, samples were dissolved into very less amount of milliQ (~5μl) to increase their apparent concentration as our reaction yields are relatively low. These were given for MALDI (Matrix assisted laser desorption ionization) analysis.

High resolution mass spectrometry (HRMS):

HRMS, a mass spectrometric method, is a high resolution technique in which HPLC is coupled with mass spectrometry. Using this technique, mass characterization of non-volatile, thermally labile and charged molecular species can be carried out. In HRMS, The HPLC column is directly connected to mass analyzer instead of the detector. The separated peaks are converted in to ions using ion sources and characterized by the analyzer.

HRMS characterizations of the samples were carried out in collaboration with Dr. Ajeet Singh at the CAMS Venture Center at N.C.L. Innovation Park. 1260 Infinity Nanoflow LC was used for high resolution purification and 6540 UHD Accurate Mass QTOF was used as mass analyzer.

Results and discussion:

1a. Oligomerization of AMP in the presence of DLPC and amino acids

AMP is known to polymerize under DH-RH cycling conditions in the presence and absence of lipids (Mungi and Rajamani, 2015), with the yields being higher in the former case. RNA-like short oligomers was observed in each set of reactions with varying yield. Figure 1 depicts the oligomerization of AMP in the presence of DLPC and Asp over the cycles 0, 1, 4 and 7, respectively.

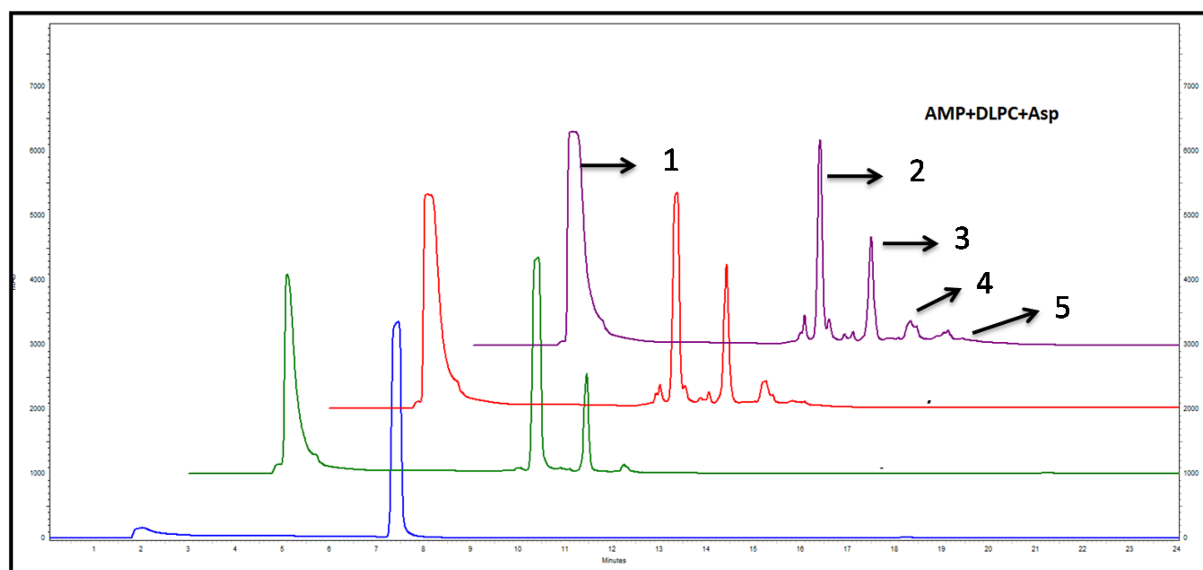


Figure 1: HPLC chromatogram of AMP oligomerization in the presence of DLPC and Asp over cycles 0, 1, 4 and 7, respectively. The indicated peaks are 1) breakdown product Adenosine 2) AMP monomer 3) AMP dimer like oligomer 4) AMP trimer like oligomer 5) AMP higher-mers.

It has previously been reported that in the presence of lipids, AMP oligomerizes to a better extent. To understand whether there is an additve role of lipids and amino acids, HPLC chromatogrms of the main reaction set was compared with the control reactions (Figure 2). Reactions with lipids showed better oligomerization than in reactions with just AMP and AMP+Asp. AMP qualitatively showed somewhat similar oligomerization with and without Asp.

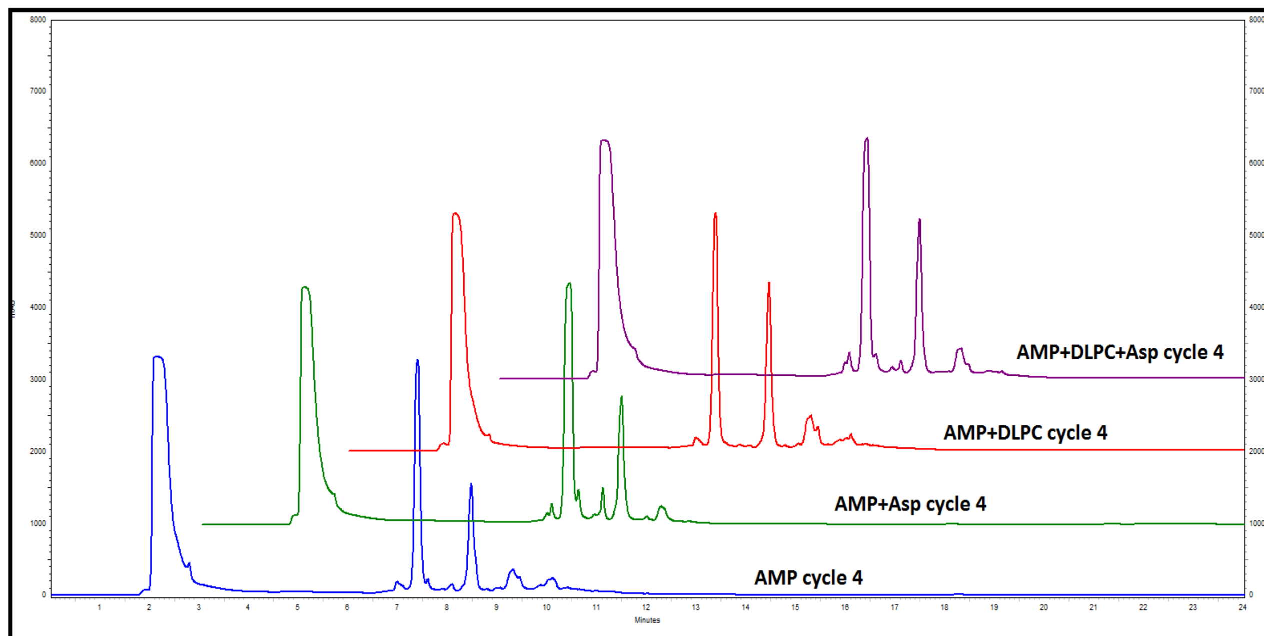
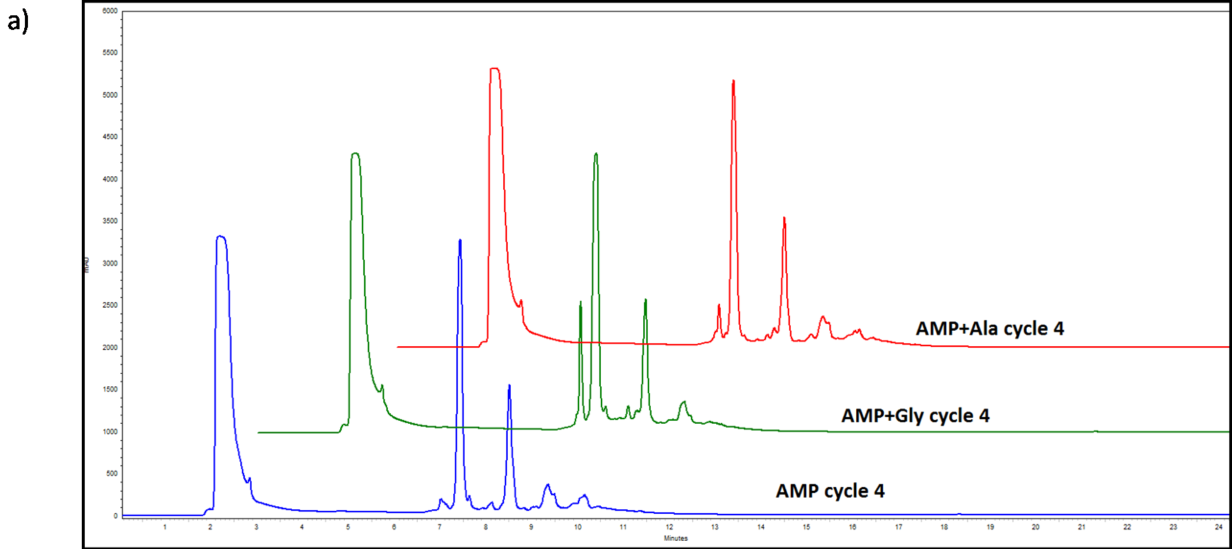


Figure 2: HPLC chromatogram comparison of main reaction (AMP with DLPC and Asp) along with the control reactions.

Additionally, in the presence of Gly, Ala and Asp too, AMP showed somewhat similar oligomerization activity as was seen in the AMP control reaction. AMP in the presence of Val showed slightly better dimer formation. All these are indicated in Figures 3a and 3b.

In the presence of both DLPC and amino acids, AMP qualitatively showed better oligomerization in the presence of Asp and Ala among GADV (Figure 4). Reaction with both DLPC and Val seemed to work the least as the products were lesser than what was seen in that of AMP control reaction. The reactions containing lipid and Gly/Ala/Asp (Figure 4) showed better oligomerization in the presence of lipid in comparison to when they were present with only AMP (Figure 3) . Therfore, GADV amino acids seem to have had no significant role on the oligomerization of AMP under DH-RH conditions.

a)



b)

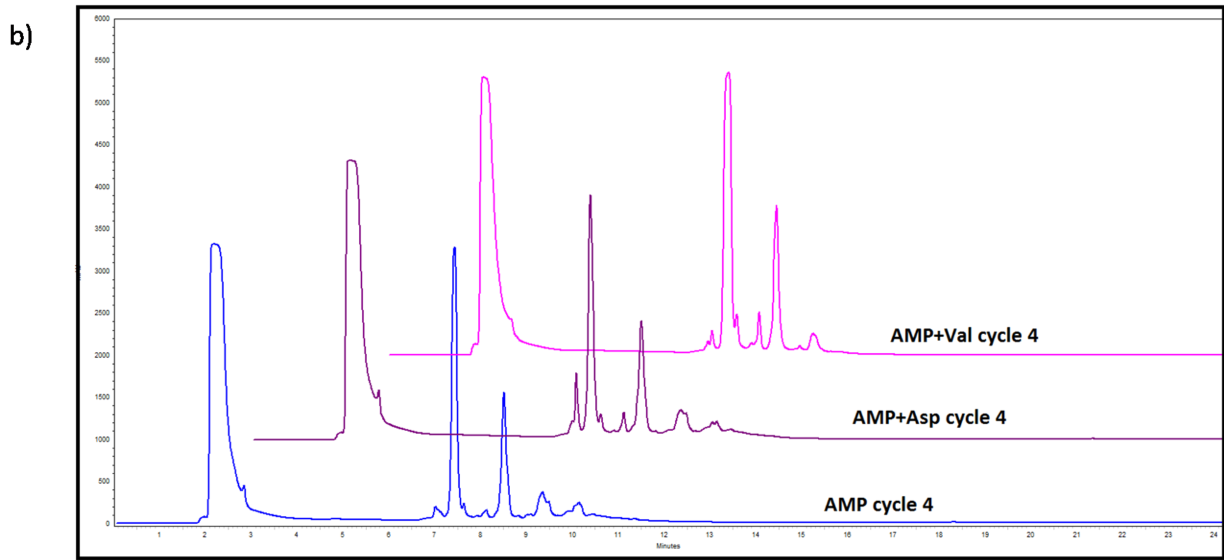


Figure 3: HPLC chromatogram comparison of AMP oligomerization in the presence of Gly, Ala, Asp and Val.

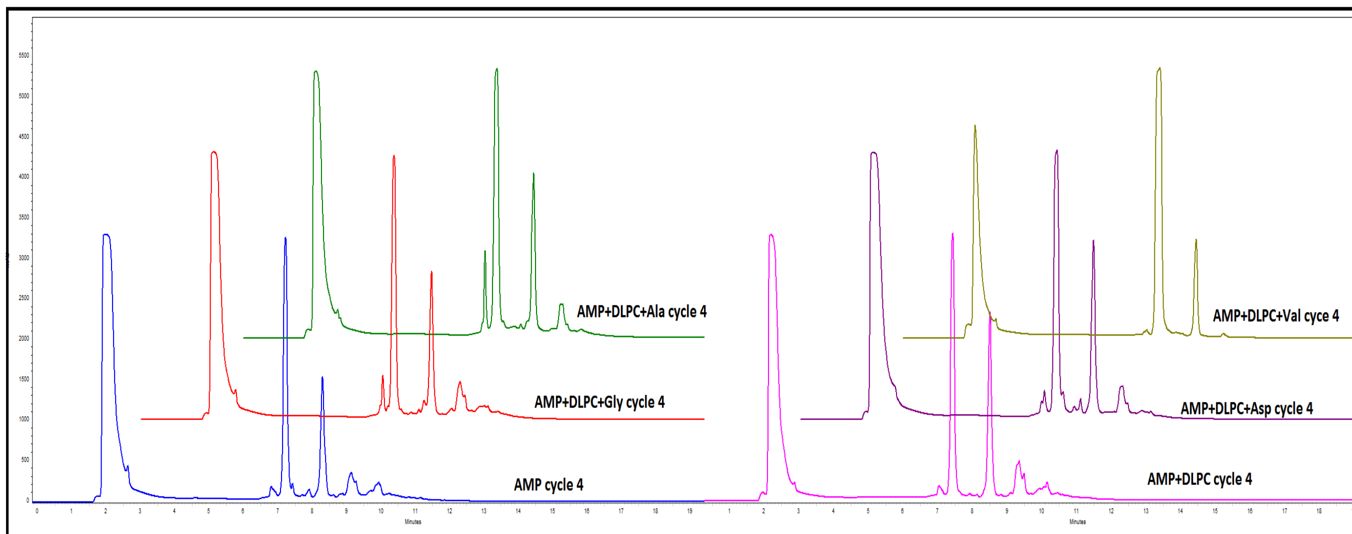


Figure 4: HPLC chromatogram comparison of AMP oligomerization in the presence of both lipid plus various amino acids.

1b. Depurination in DH-RH cycling reactions involving non-activated nucleotides:
Depurination of AMP monomer was observed soon after one DH-RH cycle. In order to understand whether the presence of lipids and/or amino acids might have any protective effect on the nucletiodes, we carried out HPLC runs with lower sample loads to get sharp clear peaks for purposes of quantitation (Figure 5). As mentioned earlier, breakdown appeared to be less in the case of reactions containig lipids, which is somewhat apparent from the HPLC run too (Figure 5).

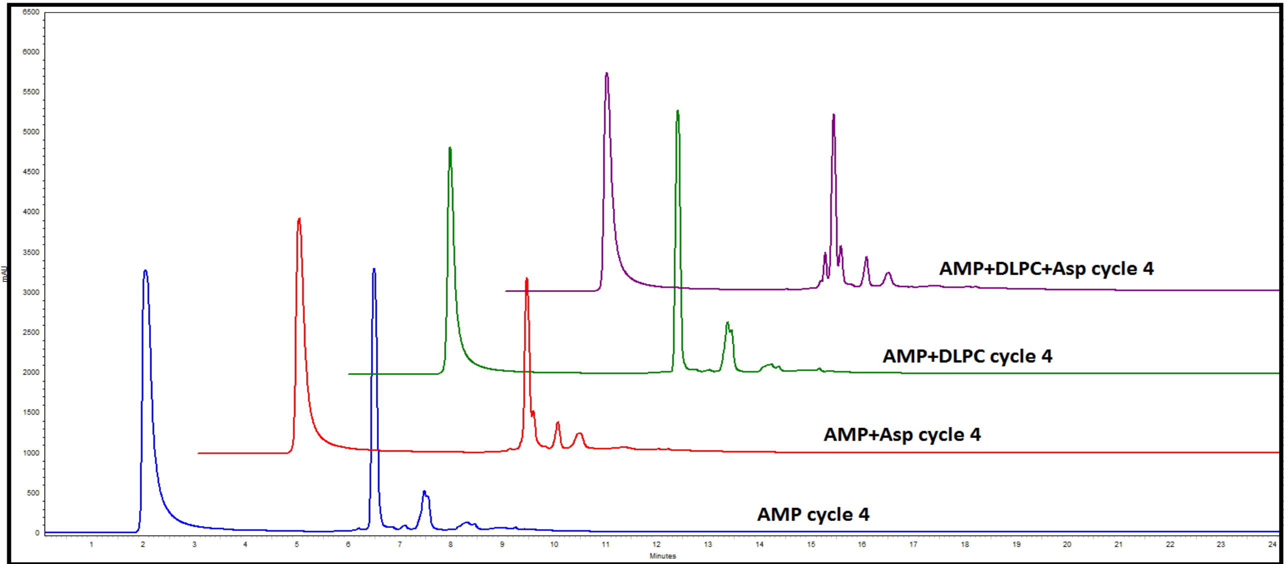


Figure 5: Representative HPLC chromatogram showing depurination during AMP oligomerization in the presence of DLPC and amino acids.

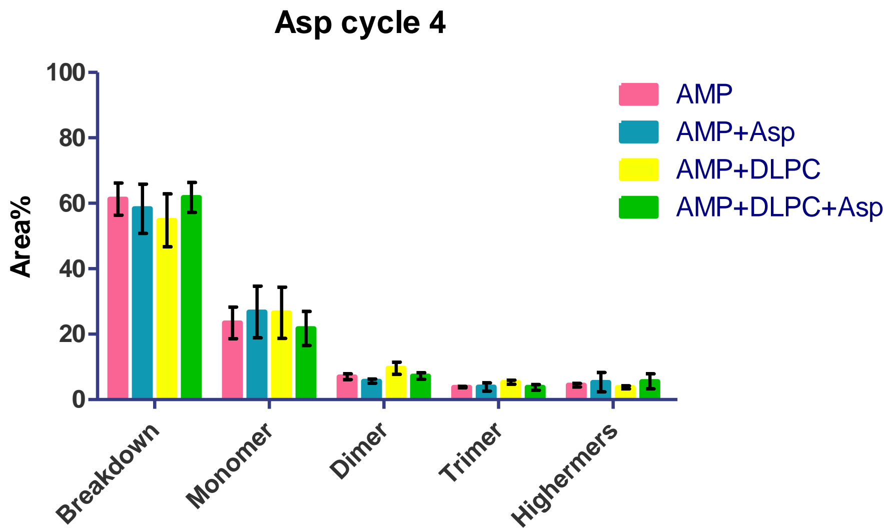


Figure 6: Quantitative area analysis of HPLC chromatograms of the control reactions and Asp reaction (cycle 4).

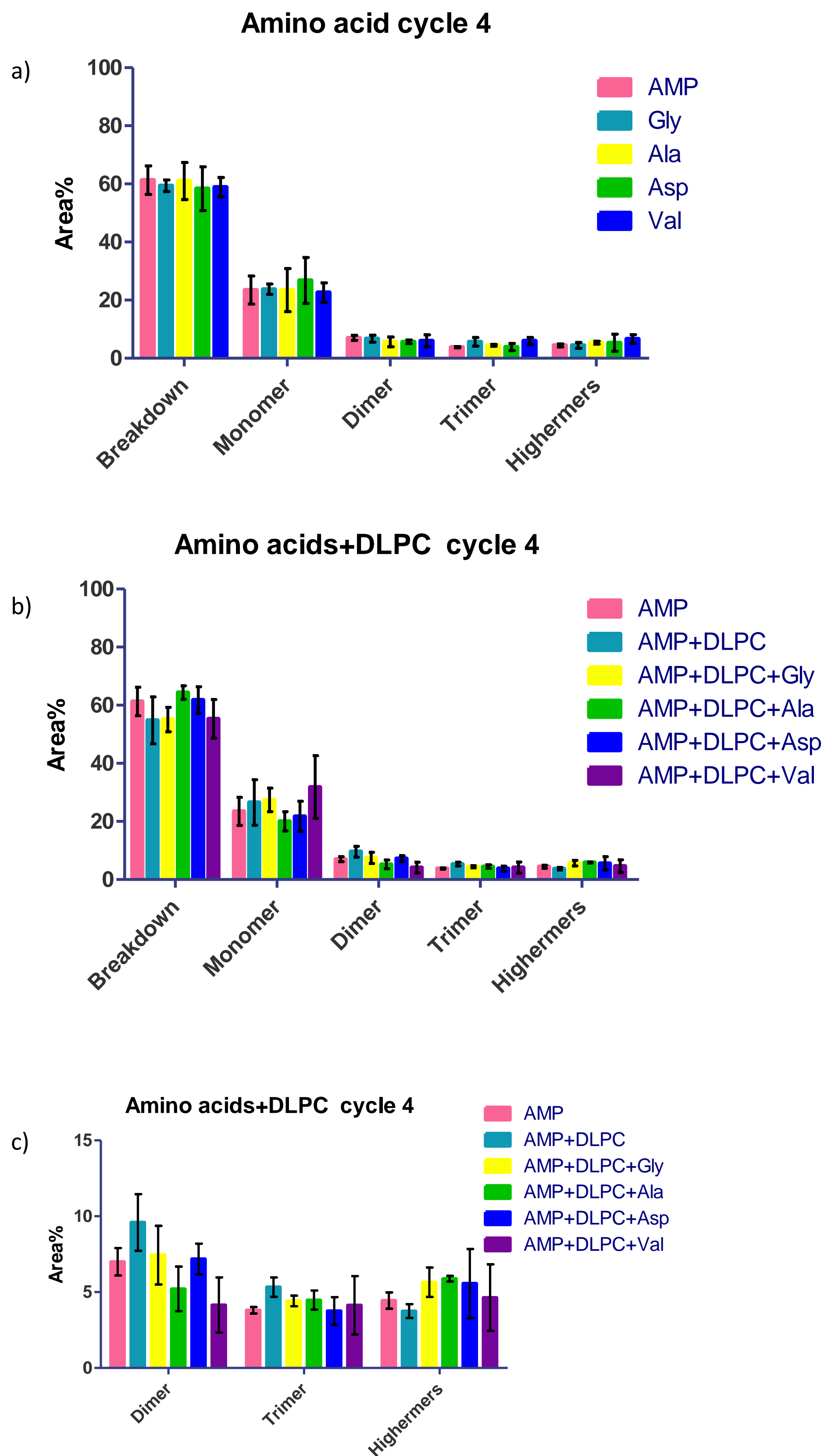


Figure 7: Comparison of quantitative area analysis of a) only amino acids b) amino acids+ lipid c) Zoomed view of oligomers in 7b.

Quantittavive area% analysis of reactions (which were carried out in triplicates) showed that around 60% of the monomers got depurinated after four cycles. No decrease in the % of depurination was observed in the presence of both Asp and lipid in comaprison to that of the AMP only control reaction (Figure 6). The presence of lipids, in general, conferred some protection from depurination which is, however, not apparent in the Asp-based reaction due to large error bars. The presence of amino acids did not confer any significant protection from depurination in the monomers, either, alone or in the presecne of lipids (Figure 7). Even though DLPC seemed to favour formation of more dimers, when amino acids were also present in the miliue, the mixture of lipid and amino acid seemed to slighltly facilatate the highermers over dimers.

The HPLC analysis that we carried out is only indicative of what is probably happening in the non-activated nucleotide based reactions from a qualitative standpoint. In order to be more precise and, potentially be able to quantify these reactions, mass characterization of the samples were carried out. MALDI could pick the mass for AMP depurinated dimer and trimer as shown in Table 5 and Figure 8. However, many of the time point samples did not fly well on MALDI. The reason for this could be the presence of sodium salt in the reaction mixtures as sodium salt of AMP was used as the starting material in the reactions. To get rid of the salt, both, dialysis using a membrane and HPLC using C-18 column, had been carried out. However the MALDI intensities did not improve considerably.

Sl/No	Sample	Expected mass	Observed mass
1	AMP(control)	347.0631	348.0352
2	AMP+ Na^+	370.0523	370.0183
3	Dimer-1base	559.0717	560.0266
4	(Dimer-1base)+Na^+	582.0614	582.0075
5	Trimer-2base	771.0803	772.0941
6	(Trimer-2base)+Na^+	794.07	794.0262

Table 5: Mass obtained from MALDI from Control and reaction samples.

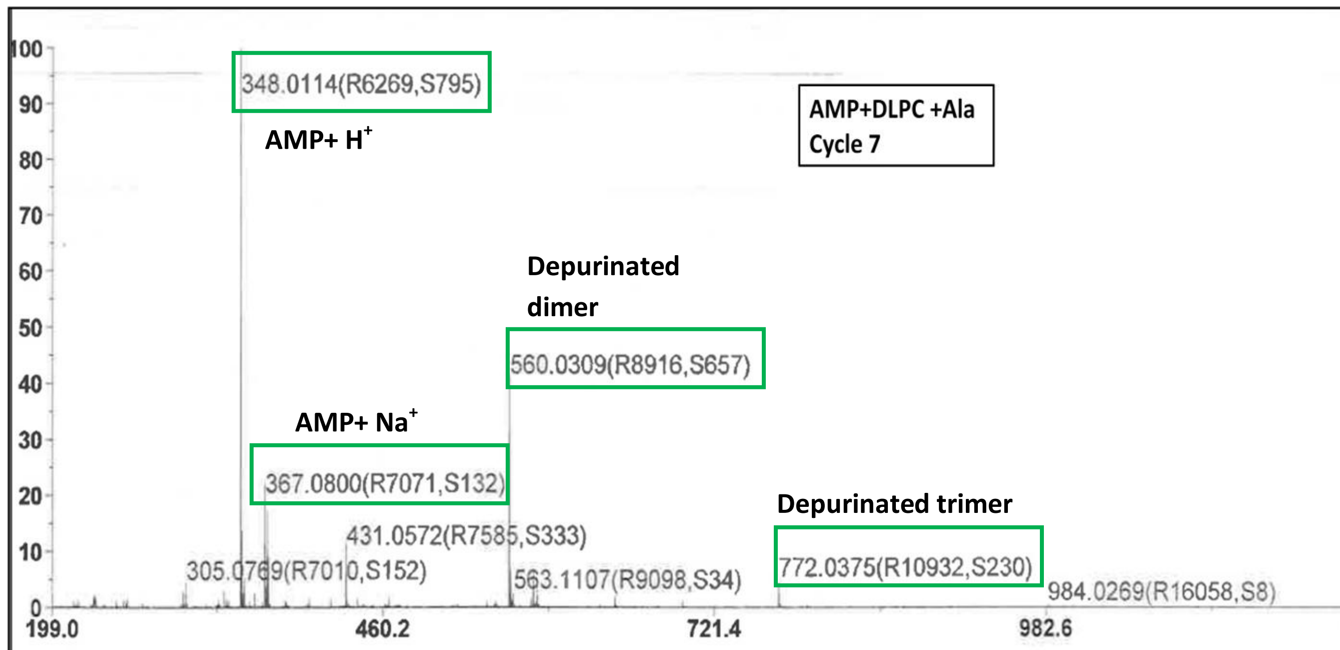
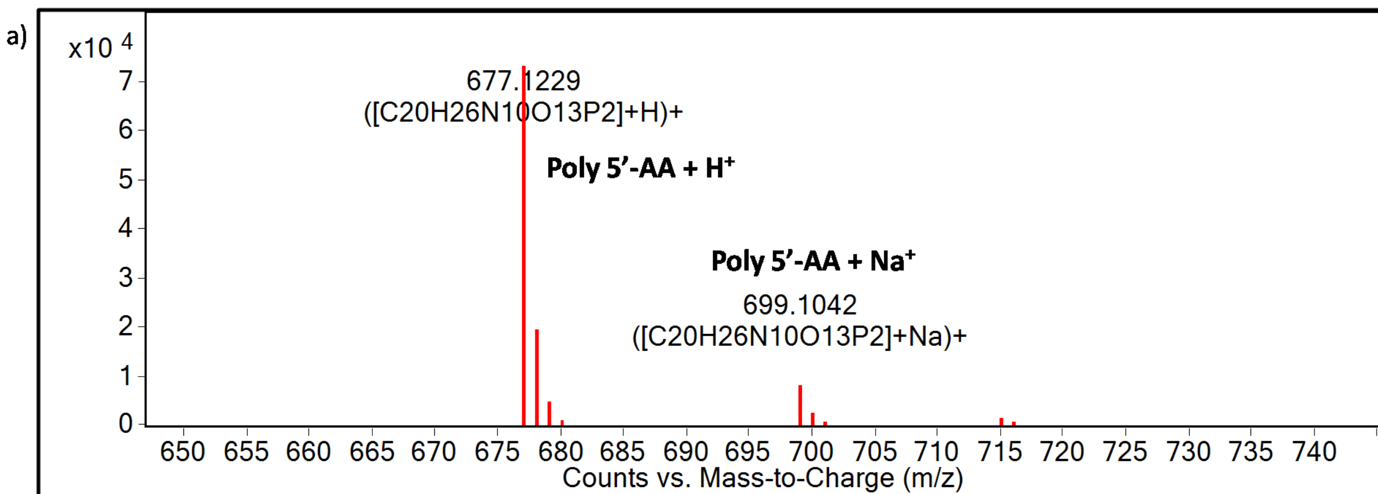


Figure 8: MALDI spectrum representing peaks from the reaction containing AMP+DLPC+Ala cycle 7.

In order to get better mass spectrum of the resultant oligomers in our reactions, samples were subsequently analyzed using HRMS. Standardization of control samples including AMP, Poly 5'-AA dimer, Poly 5'-AAA trimer (Figure 9a and 9b) has been succefully completed (Table 6). The HRMS analysis of the reaction samples are ongoing and will be completed shortly.



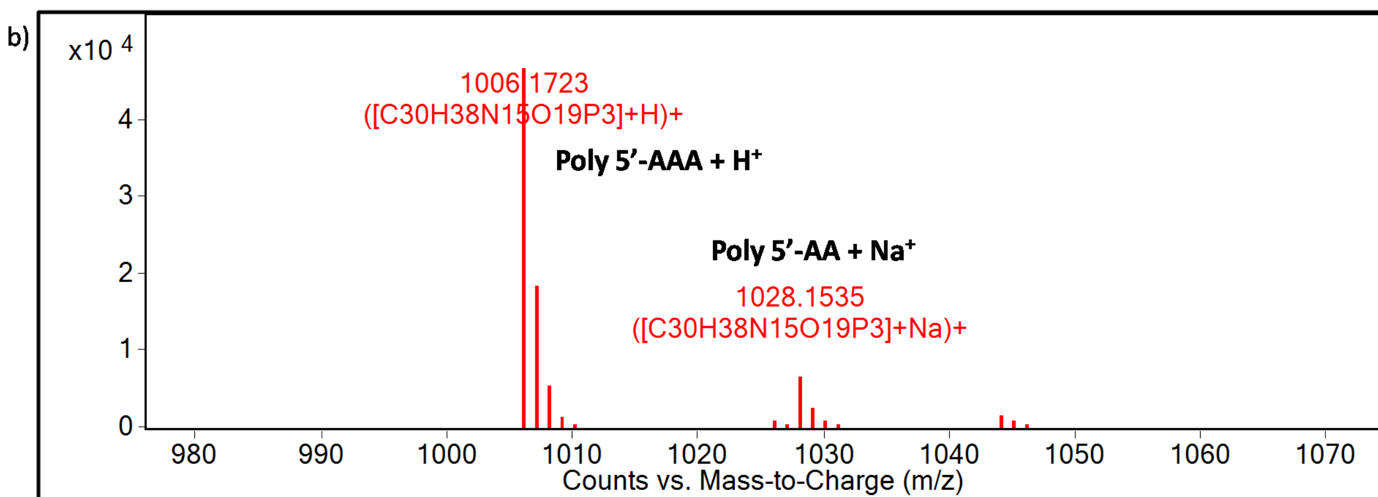


Figure 9: HRMS standardization for a) AA dimer b) AAA trimer

Sl no	Sample	Reference mass	Observed mass
1	Poly 5-AA	676.1156	676.1155
2	Poly 5-AAA	1005.168	1005.168
3	ImpG	413.0849	413.0853
4	Guanosine 5'-monophosphate (GMP)	363.058	363.0584
5	Poly G_4	1318.234	1318.235
6	Poly G_5	1663.281	1663.287

Table 6: Standardization of control samples for non activated and activated nucleotide reactions using HRMS.

2. **Activated nucleotide oligomerization under different prebiotic conditions:**

a) **Role of DLPC and alanine as co-solutes on the oligomerization of ImpG at room temperature.**

No oligomerization of ImpG was observed in the presence of DLPC and Ala (Figure 10). ImpG got hydrolyzed to GMP after 5days of reaction. Not surprisingly, DLPC conferred protection on the activated nucleotide from losing its imidazole ring. However, very interestingly, in the presence of Ala, a chemical species absorbing at 260nm eluted without interacting with the column. Additionally, in the two reactions involving Ala, a reduced peak for GMP was observed (Figure 10). It is plausible that the positively charged amino group of Ala could interact with the negatively charged phosphate of GMP, resulting in a complex which is neutral in charge, which could potentially elute in the dead volume. Mass characterization of this particular peak is ongoing, which should clarify the real nature of this peak. Standardization of control samples for HRMS including that of Poly G_4 and Poly G_5, has been successfully completed (Figure 11a and 11b).

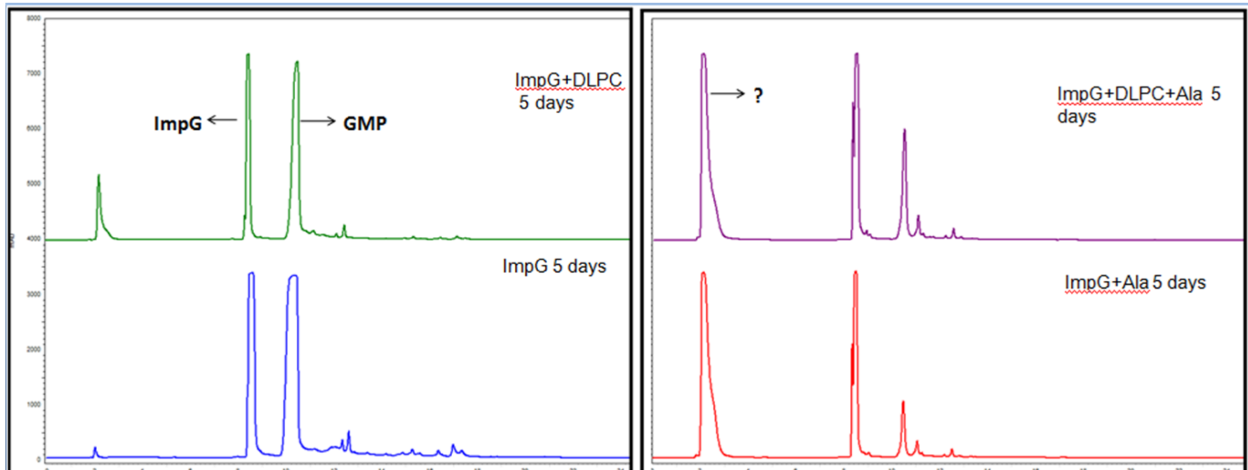
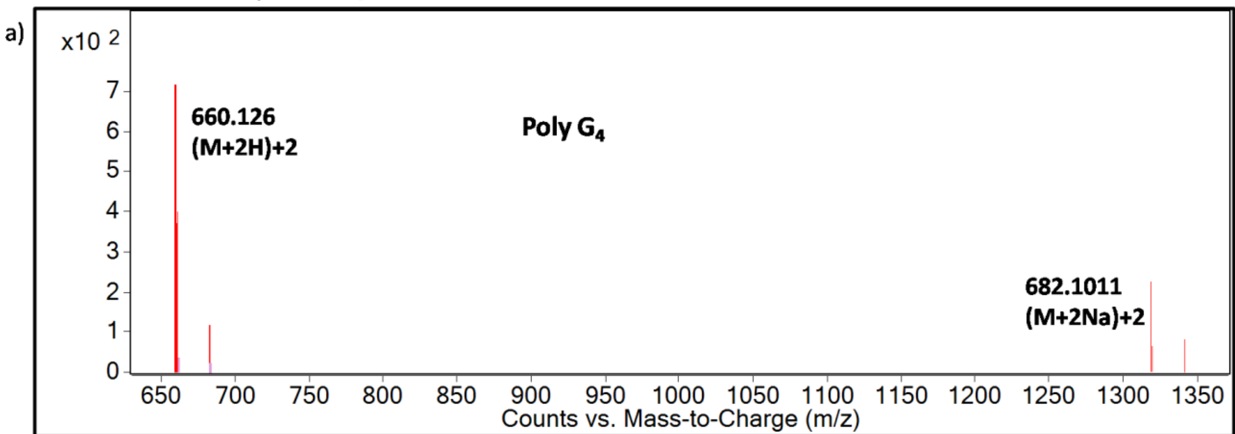


Figure 10: HPLC chromatograms of ImpG, ImpG+ DLPC, ImpG+ Ala and ImpG+ DLPC+ Ala 5th day time point.



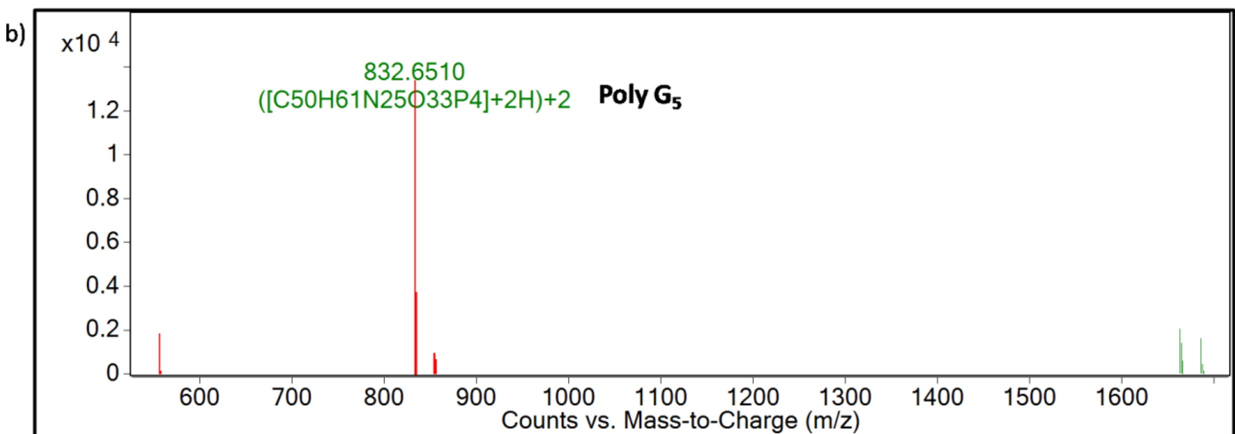


Figure 11: HRMS standardization of a) Poly G_4 b) Poly G_5

b) Dehydration reaction of ImpG in the presence of DLPC and alanine:

Preliminary HPLC runs showed no significant oligomerization that had occurred in any of the conditions. As suspected, oligomerization of activated nucleotides is possibly not as favoured at temperatures higher than room temperature as the imidazolides are known to be temperature sensitive and could easily lose the leaving group (imidazole). The temperature in these reactions was around 30ºC.

c) Oligomerization of ImpG in the presence of DMPC/POPC under vacuum at room temperature:

Reaction samples got dehydrated partially in the first few hours itself and fully dried in one day. ImpG got hydrolyzed to GMP after 5 days in reactions, both, with and without lipids to similar extent (Figure 12). Two unknown peaks were clearly observed after 5 days. One peak eluted prior to ImpG and other eluted after GMP. The mass characterizations of both these peaks are ongoing.

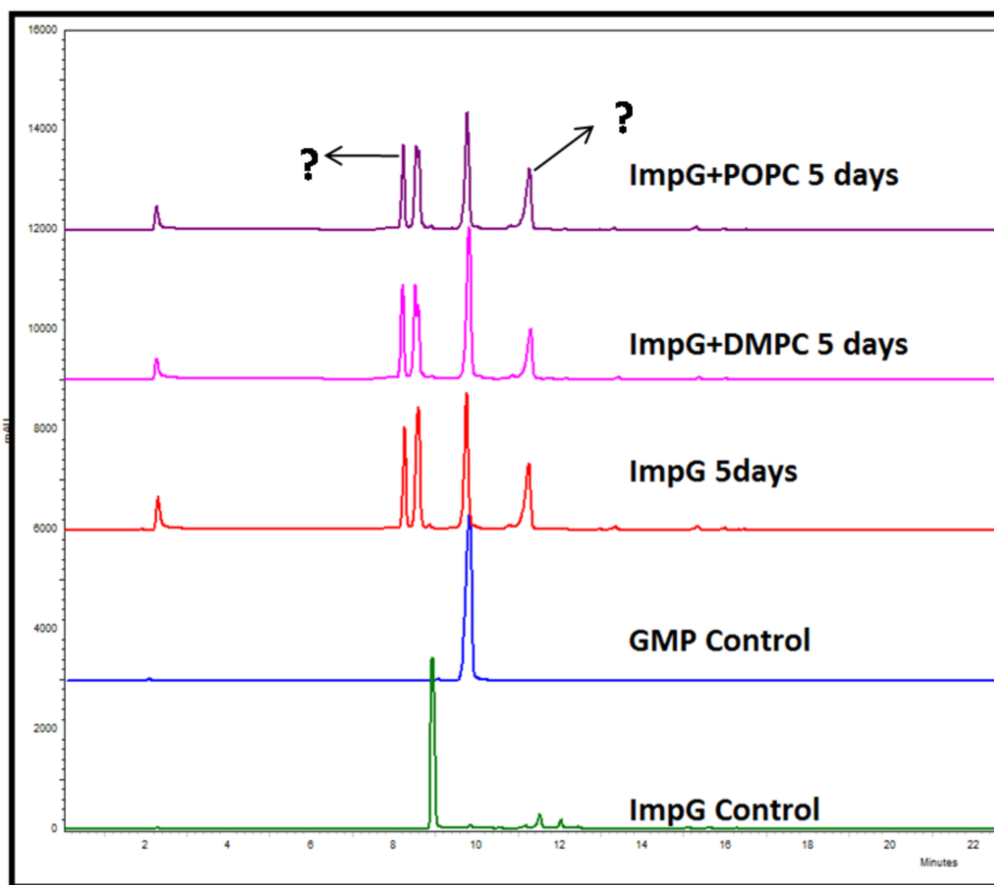


Figure 12: Comparison of ImpG, ImpG+ DMPC and ImpG+ POPC 5th day time point with ImpG and GMP control.

d) Oligomerization of ImpG in the presence of DMPC at room temperature:

This reaction was inspired by the X-ray data generated in Dr. Raghunathan's lab at RRI. Even though there was no obvious ImpG oligomerization, the imidazole activation seemed to be sustained in a higher ratio for a longer period in the presence of DMPC (Figure 13) as compared to the DLPC based reactions discussed earlier (Figure 10). This could, however, be reflective of the fact that the DMPC containing reaction had a higher concentration of lipid (12.5mM) in comparison to the DLPC reaction where its concentration was only 1mM. Therefore, to understand if a certain lipid conferred greater stability to the activated monomer in a prebiotic context, reactions with comparable parameters have to be carried out.

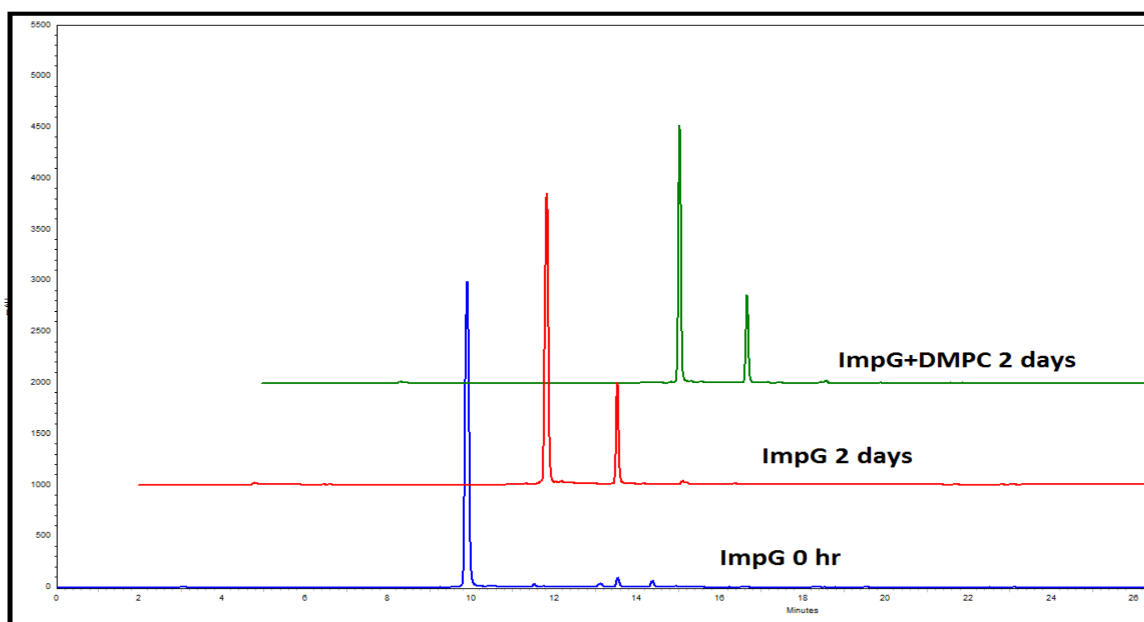


Figure 13: HPLC chromatogram of 2nd day time point of ImpG and ImpG+ DMPC.

In conclusion, the amino acids studied including Gly, Ala, Asp and Val did not seem to enhance oligomerization of either the non-activated or activated nucleotides.

Furthermore, they also did not have any noticeable synergistic on the oligomerization of AMP under DH-RH cycling conditions, effect in the presence of DLPC.

As reported in previous studies, presence of lipid did confer some protection for AMP monomers from depurination in the non-activated nucleotide based studies. However, the presence of amino acids did not provide any significant protection of AMP either, alone or in the presence of lipids. Finally, ongoing studies are mainly centered on mass characterization of resultant products from many of the aforementioned reactions. Systematic characterization will hopefully enable tying up some of the incomplete observations that have been reported.

References:

An, R., Jia, Y., Wan, B., Zhang, Y., Dong, P., Li, J., and Liang, X. (2014). Non-enzymatic depurination of nucleic acids: factors and mechanisms. PLoS One *9*, e115950.

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

Benner, S.A., Kim, H.J., and Carrigan, M.A. (2012). Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. Acc Chem Res *45*, 2025-2034.

Ferris, J.P., and Ertem, G. (1993). Montmorillonite catalysis of RNA oligomer formation in aqueous solution. A model for the prebiotic formation of RNA. J Am Chem Soc *115*, 12270-12275.

Gilbert, W. (1986). Origin of life: The RNA world. Nature *319*, 618-618.

Huang, W., and Ferris, J.P. (2003). Synthesis of 35-40 mers of RNA oligomers from unblocked monomers. A simple approach to the RNA world. Chem Commun (Camb), 1458-1459.

Ikehara, K. (2014). [GADV]-protein world hypothesis on the origin of life. Orig Life Evol Biosph *44*, 299-302.

Kanavarioti, A., Monnard, P.A., and Deamer, D.W. (2001). Eutectic phases in ice facilitate nonenzymatic nucleic acid synthesis. Astrobiology *1*, 271-281.

Mamajanov, I., MacDonald, P.J., Ying, J., Duncanson, D.M., Dowdy, G.R., Walker, C.A., Engelhart, A.E., Fernández, F.M., Grover, M.A., Hud, N.V.*, et al.* (2014). Ester Formation and Hydrolysis during Wet–Dry Cycles: Generation of Far-from-Equilibrium Polymers in a Model Prebiotic Reaction. Macromolecules *47*, 1334-1343.

Miller, S.L. (1953). A Production of Amino Acids Under Possible Primitive Earth Conditions. Science *117*, 528-529.

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

Patel, B.H., Percivalle, C., Ritson, D.J., Duffy, C.D., and Sutherland, J.D. (2015). Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. Nat Chem *7*, 301-307.

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

Powner, M.W., Sutherland, J.D., and Szostak, J.W. (2010). Chemoselective multicomponent one-pot assembly of purine precursors in water. J Am Chem Soc *132*, 16677-16688.

Rajamani, S., Vlassov, A., Benner, S., Coombs, A., Olasagasti, F., and Deamer, D. (2008). Lipid-assisted synthesis of RNA-like polymers from mononucleotides. Orig Life Evol Biosph *38*, 57-74.

Saladino, R., Carota, E., Botta, G., Kapralov, M., Timoshenko, G.N., Rozanov, A.Y., Krasavin, E., and Di Mauro, E. (2015). Meteorite-catalyzed syntheses of nucleosides and of other prebiotic compounds from formamide under proton irradiation. Proc Natl Acad Sci U S A *112*, E2746-2755.

Stark, B.C., Kole, R., Bowman, E.J., and Altman, S. (1978). Ribonuclease P: an enzyme with an essential RNA component. Proceedings of the National Academy of Sciences *75*, 3717-3721.

Weber, A.L. (1989). Thermal synthesis and hydrolysis of polyglyceric acid. Orig Life Evol Biosph *19*, 7-19.

Wieczorek, R., Dorr, M., Chotera, A., Luisi, P.L., and Monnard, P.A. (2013). Formation of RNA phosphodiester bond by histidine-containing dipeptides. Chembiochem *14*, 217-223.

Zaug, A.J., and Cech, T.R. (1980). In vitro splicing of the ribosomal RNA precursor in nuclei of Tetrahymena. Cell *19*, 331-338.