UNIVERSITY OF GRONINGEN

CHEMISTRY BACHELOR THESIS

Complex Coacervates as a Pre-Biotic Compartment for RNA Oligomerisation

Author:
Victor B. VERDUIJN*

**s3109364*
v.b.verduijn@student.rug.nl

Assessors:
prof. dr. Sijbren OTTO
prof. dr. Johannes E. M. N. KLEIN

Daily Supervisor:
Marcel J. ELEVELD, MSc

First draft: 28th of June 2019
Final version: 4th of July 2019

rijksuniversiteit
groningen

Abstract

The origin of life is a fundamental question that has kept humanity busy since the dawn of its existence. It should, therefore, come as no surprise that there is no lack of theories which set out to answer or shed light on this intriguing question. One such theory is the RNA world hypothesis. This theory suggests that RNA predates DNA and proteins, and bases its argument on the catalytic and information storage capabilities of this macromolecule. However, this raises the question as to how the first RNA oligomers and/or polymers were formed. It was researched whether a liquid-liquid phase separation phenomenon referred to as complex coacervates could hold the key to answering this question. Complex coacervates are spherical droplets which strongly partition solutes, and could, therefore, have acted as pre-biotic compartments for primordial RNA oligomerisation. The results were analysed by means of ultra performance liquid chromatography, mass spectrometry, and NMR spectroscopy. Complex coacervate systems were identified that successfully partition RNA monomers, and their pre-activated counterparts, phosphorimidazolides. These compounds (adenosine 5'-monophosphorimidazolide and adenosine 5'-monophosphor-(2-methyl)-imidazolide) were successfully synthesised in up to 80% yields. However, phosphoramidates and guanosine phosphorimidazolides, different activated RNA monomers, could not be synthesised successfully. Even though complex coacervates can partition phosphorimidazolides no oligomerisation products had been observed. Therefore, RNA monomers were attempted to be activated in situ with poly-L-histidine, an imidazole moiety containing macromolecule, and EDC. This method was proposed to allow activation similar to templated RNA oligomerisation reactions. However, mostly hydrolysis, pyrophosphate, and buffer-nucleobase adduct formation occurred without any oligomerisation product. Therefore, it was concluded that complex coacervates can efficiently partition RNA monomers, and to a limited extent their activated counterparts. However, more research is necessary to ascertain the feasibility of RNA oligomerisation in complex coacervates.

Contents

1 Introduction

1.1 Origins of Life: the RNA World Hypothesis

Humanity has long sought for answers concerning the origins of life. It should, therefore, come as no surprise that theories regarding this fundamental question are abundant in folklore, religion and science. Moreover, this question is especially relevant from a chemical perspective. How did modern life, with all its complexity such as DNA, proteins, RNA, lipid membranes, neurons, etc. evolve from simple chemicals available in pre-biotic history. In an attempt to answer this question, one should consider: What constitutes life? Even though there is no formal definition of life, it is generally agreed upon that life should consist of at least the following three elements: self-replication (the ability to generate (genetic) copies of itself), metabolism (the ability to convert other chemicals into useful chemicals and energy for survival), and compartmentalisation (there should be well-defined boundaries to the living system, for multiple reasons). Modern organisms all fulfil these properties of life, they have compartmentalisation due to their cell-membranes, metabolism, since they convert food into useful sugars, energy and other essentials, and they are capable of reproduction.

The ability to reproduce is of great importance to the sustenance of life's information. Without this property, the information has no means of transcending thermodynamic equilibrium (death on the molecular level) and is, thereby, lost. This raises the question: How do modern life forms circumvent the problem of information loss at equilibrium?

All known living organisms rely on self-replication or reproduction by means of DNA replication. DNA is a long polymer consisting of deoxyribonucleotides (dNMP's), of which there are 4 variants (A, C, G, T). This compound is the macromolecule responsible for the preservation of genetic information and its sequence codes for specific proteins which allow the organism to metabolise, compartmentalise and ultimately to reproduce by creating new copies of itself. However, a key piece of information is missing from the puzzle. How did DNA form if the proteins necessary for replication originate from DNA? This is where RNA comes in. This macromolecule can, like DNA, convey genetic information by making use of analogous ribonucleotides consisting of four variants (more on this in **section 1.1.1**). Moreover, it also possesses the capability of catalysis[1,2], and it is found prevalently in ribosomes[3]. These macromolecules are responsible for the translation of mRNA (messenger RNA from transcribed DNA) into proteins. Secondly, it has been demonstrated that RNA oligomers and monomers can be synthesised pre-biotically and oligomerisation can occur relatively enantio-selective by making use of montmorillonite clay[4–6]. Notable, however, are the low yields of the reaction due to the large amount of side reactions which can occur. For instance, pyrophosphate formation, and adducts of these compounds.

With the catalytic and information storage capabilities of RNA in mind, the RNA world hypothesis states that the macromolecule RNA predates DNA and proteins. It is proposed that before DNA and proteins, life was governed by RNA. However, this raises the question of how the first functional RNA polymers/oligomers were formed.

1.1.1 Monomers

As briefly mentioned before RNA is a macromolecule consisting of ribonucleotides with four different nucleobases, namely, A (adenine), C (cytosine), G (guanine), and U (uracil) (figure 1). In nature, these ribonucleotides or nucleotides for short are selectively linked by means of 3'-5' linkages. This means that the phosphate is connected to the 3' position on one ribose and on the 5' position on the other. However, upon random combination of nucleotides, other linkages appear as well, e.g. 2'-5'. This makes the abiotic oligomerisation of NMP's difficult due to the large possibility for structural isomers (figure 2)[7,8].

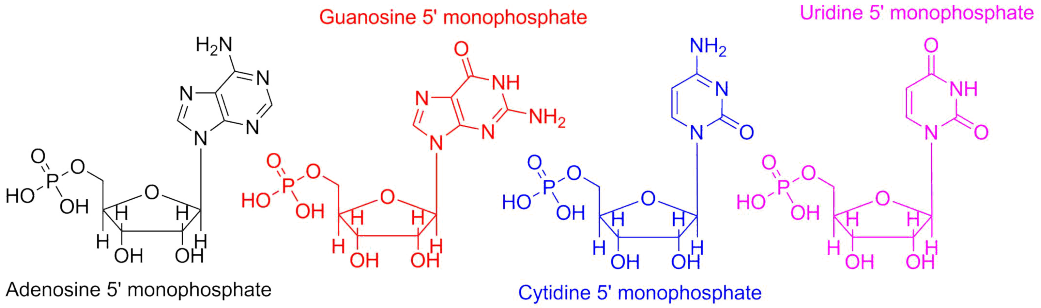


Figure 1: Chemical structure of AMP, GMP, CMP, and UMP respectively. Demonstrating the ribose monophosphate backbone connected to the nucleobase on the 1' position of ribose.

RNA, just like DNA, carries genetic information by differing the sequence of nucleotides. A single strand of RNA can be replicated in two steps. Firstly, the complementary strand is formed by means of canonical base-pairs (often with the help of proteins) and condensation reactions leading to phosphodiester links. Canonical base-pairing is the pairing of nucleotides which complement each other's hydrogen bonds, namely, adenosine with uridine, and guanosine with cytidine. Secondly, once the complementary strand has formed, the strands need to be separated such that a copy of the original template strand can be formed from the complementary strand (often with the help of proteins). Note, that both the complementary strand and the template strand carry the same genetic information while having different sequences.

1.1.2 Oligomerisation Methodologies

Now that the monomeric nature of RNA has been discussed, the transition is made towards the oligomerisation of RNA. This process starts from NMP in solution or cellular medium, and by means of condensation reactions, the phosphodiester linkages are formed. This can, as previously touched upon, occur in different ways (figure 2).

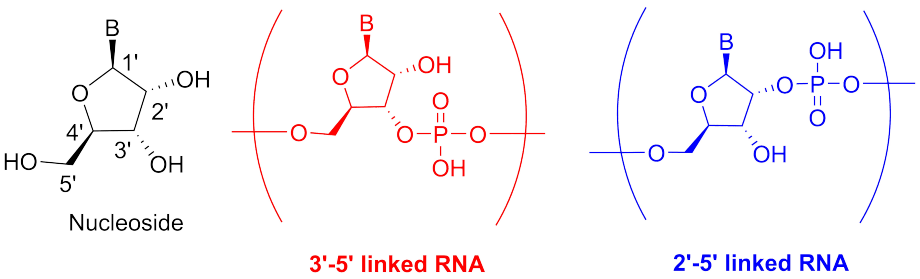


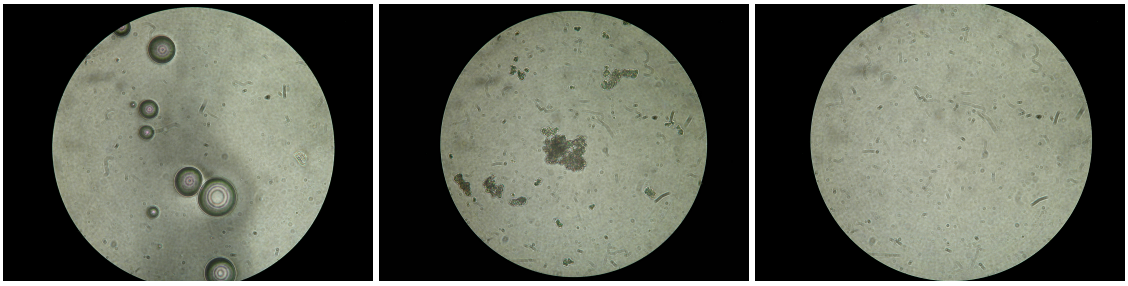
Figure 2: Left to right: **1).** Chemical structure of nucleoside with nucleobase B demonstrating the position nomenclature for nucleosides. **2).** 3'-5' phosphate linkage in RNA. **3).** 2'-5' phosphate linkage in RNA. N.B. random mixtures of phosphate linkages in the macromolecule is possible.

Moreover, the prevalent type of linkage can be influenced by making use of a catalyst[2,4]. Natural proteins, for instance, produce the 3'-5' linked variant, whereas some abiotic catalysts result in random mixtures of different linkages[8,9]. Moreover, catalysts are required in order for this reaction to occur spontaneously in quantitative amounts due to the inherent competition of oligomerisation with hydrolysis and the stability of the monomers[10]. Therefore, different activation methodologies have been suggested which increase the reactivity of the phosphate group and/or aid the monomers in obtaining the right orientation to react. Examples of these activation methodologies are: phosphorimidazolides[6], OAt-ester activated phosphate[11], phosphoramidates[4,12], complementary strand templating reactions[2,11], montmorillonite clay[6,13] and cyclic NMP's[8,9] or combinations thereof.

The specifics of the techniques used during the research project (phosphoramidates, cyclic NMP's and phosphorimidazolides) will be discussed in **section 2**. However, what montmorillonite clay, eutectic ice phases, and templated reactions have in common is the high local concentration of monomers that is required for the oligomerisation to occur competitively with hydrolysis. Which brings us to **section 1.2**.

1.2 Complex Coacervates

A recently (1929) recognised phenomenon related to high local concentrations is the property of liquid-liquid phase-separation referred to as complex coacervation. This phenomenon can occur spontaneously when either a poly-electrolyte/macro-ion is mixed with an oppositely charged solute or poly-electrolyte in aqueous solution under the right conditions. Instead of precipitating or dissolving, the poly-electrolyte(s) form(s) highly condensed phase-separated droplets[14] which strongly partition solutes. This phase is referred to as the (complex) coacervate phase[15,16]. However, it is notable that not all poly-electrolyte-solute and complementary poly-electrolyte systems exhibit this phase separation behaviour. System properties, such as charge-density, chirality[17], temperature, pH, macro-ion concentration, and salt concentration all play a role in determining the behaviour of the mixtures of these components[17–21]. However, contrary to the many variables that determine the feasibility of complex coacervation, complex coacervation is by no means a rare phenomenon. Consequently, the identification of systems exhibiting complex coacervation with novel behaviour is all but absent from literature[15,22]. Moreover, complex coacervation also presents itself in nature. As an example, it was recently demonstrated that complex coacervation plays a role in the adhesive that mussels use to attach to wet surfaces[23]. These systems have in common that the different system parameters all influence the relative interaction strength between the components, and a delicate balance has to be obtained in order for this phenomenon to present itself.



(a) Complex Coacervates **(b)** Precipitate Complex **(c)** Homogeneous Solution

Figure 3: Optical micro-graphs of different behaviours for mixtures of charge-complementary macro-ions. **3a).** Complex coacervates of poly(4-styrenesulphonate) (PSS):poly(diallyldimethylammonium chloride) (PDADMAC) (1:1 monomer ratio, 10 mM monomer concentration with 800 mM KBr in water at pH 7) **3b).** Precipitate complexes of Poly-L-Histidine:PSS (1:1 monomer ratio, 10 mM monomer concentration in water at pH 5.6). **3c).** Homogeneous solution of Poly-L-Lysine:AMP (1:3 monomer ratio, 15 mM AMP concentration in water at pH 6). (Circle is 1.5mm in diameter).

Mixtures of charge-complementary poly-electrolytes or mixtures of a poly-electrolyte with a solute can result in either the formation of a precipitate complex, complex coacervates, or homogeneous solutions (figure 3). In order to understand how each different system parameter influences the behaviour of a mixture, some understanding of the stability and formation of complex coacervation is necessary.

1.2.1 Stability and Formation

In order for complex coacervation to occur spontaneously, its change in Gibbs free energy of coacervation $\Delta G_{coac.}$ has to be negative, which holds for any spontaneous process obeying the laws of thermodynamics. Following from the definition of Gibbs free energy, or the definition of change in Gibbs free energy at constant temperature and pressure (eq. 1), complex coacervation can either be entropy driven ($T\Delta S > \Delta H$; $\Delta H > 0$), enthalpy driven ($\Delta H < T\Delta S$; $T\Delta S < 0$) or both ($T\Delta S > 0 > \Delta H$).

$$\begin{aligned} G &= H - TS \\ \Delta G &= \Delta H - T\Delta S \end{aligned} \tag{1}$$

System properties such as pH, temperature, concentration, ionic strength, chirality and charge-density influence the balance between enthalpy and entropy in the system, and hence the change in Gibbs free energy of coacervation will change accordingly; resulting in either complex coacervation, or a different mixture-behaviour (fig. 3). In general, for poly-electrolytes, complex coacervation is entropy driven; however, enthalpy can also play an important role[21]. Effects such as electrostatic interactions, counter-ion release to the bulk solution, the release of water in the hydration shell to the bulk solution, hydrogen-bonding and hydrophobicity determine the degree to which the system is capable of forming complex coacervates. In order to determine how system parameters should be tuned in order to sustain or achieve complex coacervation, the influences of different system parameters are evaluated separately.

Effect of pH on complex coacervation
Since complex coacervation is dependent on electrostatic interactions and high charge complementarity between two macro-ions, pH has a large effect on the magnitude of this complementarity by directly influencing the degree of ionisation of functional groups. Therefore, the pH of the mixture directly influences the enthalpy of coacervation by affecting the electrostatic interactions between macro-ions. Moreover, it was found that the pH of macro-ions before mixing is of importance to the stability of coacervates. Adjusting the pH after mixing neutral species, sometimes resulted in precipitates or clear solutions. It is believed this is caused by poor solubility of macro-ions at a certain pH, the abrupt change in system composition, and/or the possibility of local pH differences due to the high in-homogeneity of the mixture. Moreover, from DLVO (Derjaguin Landau Verwey Overbeek) theory, a theory about the complexation of colloidal particles, it is argued that precipitation/sedimentation/coagulation is an irreversible process[24]. Thereby, if a precipitate has been formed, reverting to the initial system conditions will not result in the re-formation of complex coacervates. To summarise, the pH should be balanced such that the highest charge complementarity of the system is obtained while still providing sufficient solubility of the macro-ions.

Effect of charge-density on complex coacervation

Charge-density has a large effect on the feasibility of complex-coacervation by influencing both the enthalpy and entropy of coacervation by means of the relative strength of electrostatic interactions, and by the amount of coacervate-induced releasable counter-ions and molecules in the hydration shells[21]. The property of charge-density is directly influenced by the pH for macro-ions for which holds that their functionalities' pKa are in the range of the system pH. However, it need not, necessarily, be influenced by the pH in all systems. For instance, when anionic groups with very low pKa values (e.g. sulphate group with a pKa of -3 in water) and cations of quaternary ammonium salts are used to induce complex coacervation. The former functional group is relatively insensitive to changes in pH, and the latter is largely immune to changes in pH due to the alkylated nature of the ammonium group; hence not relying on proton concentrations. Moreover, charge-density becomes of increased importance the smaller the macro-ion is. As an example, poly-L-lysine can form coacervates with ATP, ADP and AMP; however, it requires higher concentrations of ADP and AMP, respectively, than ATP in order to compensate for the loss of charged phosphate groups (lower charge-density compensated by concentration). Secondly, functionalisation of the phosphate group in AMP, forming, e.g., phosphorimidazolides, results in the loss of complex coacervation capabilities due to the loss of charge-density.

Effect of temperature on complex coacervation

Even though temperature itself has a mild effect on the stability of complex coacervation, changes in temperature do have an effect. Upon an increase of the temperature, the coacervate-phase will become denser, and smaller droplets will form. This behaviour is similar to polymer-solvent interactions with a lower critical solution temperature (LCST). Mixtures above this temperature will start to become more immiscible upon an increase of temperature[18]. Moreover, a decrease of temperature below the LCST is argued to result in full miscibility of the components at any composition.[18]

Effect of chirality on complex coacervation

Chirality has an effect on the capability of two poly-electrolytes to form complex coacervates in the case of poly-peptides. In 2012 Priftis and Tirrell demonstrated the possibility of preparing complex coacervates of poly-peptide mixtures under a variety of conditions[19]. Of these poly-peptide mixtures, combinations of poly-lysine and poly-glutamic acid were studied most prevalently. However, upon analysis of the optical purity of the supposedly all L-glutamic acid, it was found the compound was of racemic nature, thus being a random copolymer of D and L amino acids[25]. Upon further investigation, the same team in collaboration with other scientists (Perry et al.) reported in 2015 that chirality is of major importance to the capability of poly-peptide mixtures to undergo complex coacervation[17]. They reported the inability to form complex coacervates when both poly-peptides were optically pure, instead, stable precipitate complexes were formed. Only when, at least, one of the poly-peptides was racemic (D,L) would complex coacervation be observable. This behaviour was demonstrated to arise from the possibility of forming highly ordered regimes of beta-sheets when both poly-peptides were optically pure[17].

Effect of ionic strength on complex coacervation
Salt is known to directly influence the complexation behaviour of the mixture, resulting in either precipitation, complex coacervation or dissolution, depending on the salt concentration. This effect is caused by ions directly influencing the electrostatic interactions between the macro-ions by means of shielding, and by the added effect of counter-ion release upon complex coacervation. Therefore, for high charge-density mixtures, precipitation is expected without salt. However, addition of salt to obtain intermediate ionic strength will result in complex coacervates. Thereafter, addition of more salt is expected to result in a clear solution.[18,19] Moreover, the nature of the salt ions is also of importance to the behaviour of the mixture. It has been reported that salts follow the Hofmeister series for their effect on the critical salt concentration, namely, less hydrated anions resulted in more macro-ion screening compared to more hydrated anions[18].

Effect of poly-electrolyte chain-length on complex coacervation
Even though chain length is of major importance to the charge-density of mixed polymers it has a mild influence on complex coacervation[18]. Mostly, increasing chain-length increases the critical salt concentration due to the addition of more charged residues; thus requiring more salt for the charges to be effectively screened[18,19,21]. However, this effect is less pronounced than the system properties laid out above. Moreover, chain-length is of importance to smaller macro-ions. Combination of small chain-length polymers with relatively small macro-ions will be less likely to result in complex coacervation than larger chain-length polymers due to the lower amount of intermolecular interactions[21]. With the influence of these parameters laid out, one possesses the tools to tune the properties of complex coacervates to the desired properties.

1.2.2 RNA oligomerisation in complex coacervates

In order for RNA oligomerisation to occur efficiently, high concentrations of monomers are required. Monomer partitioning inside complex coacervate droplets increases the local concentrations is, therefore, believed to be a possible platform for enzyme-free, untemplated RNA oligomerisation. This combines two key properties of life, namely, replication and compartmentalisation. There are theories that complex coacervates acted as membrane-less protocells[7,26], and there is experimental evidence that complex coacervates can form from macro-ions with RNA oligomers or even monomers[15,16,27,28].

1.3 Research

Combining RNA oligomerisation and complex coacervation could be of interest in investigating the pre-biotic oligomerisation of RNA from monomers. Therefore, it was researched whether complex coacervates can act as a pre-biotic compartment for RNA oligomerization.

2 Materials and Methods

2.1 Chemicals

All poly-electrolytes, NMP's and imidazole analogues, unless otherwise stated, were obtained from Sigma Aldrich Corporation. Poly-(D,L)-glutamic acid sodium salt and poly-(D,L)-lysine hydrobromide were obtained from Alamanda polymers, Inc. GMP was obtained from Carbosynth Limited. Chemicals were used without further purification. Poly-peptides were solubilised by means of sonication.

2.2 Ultra Performance Liquid Chromatography

Samples were analysed, largely, by means of Ultra Performance Liquid Chromatography. In order to obtain a good separation, pre-made methods, by M.J. Eleveld, MSc, were used and tuned. These methods make use of a reverse phase C_{18} column with acetonitrile and water as eluent with $0.1\%_{vol.}$ formic acid (FA) as a modifier. Two methods were used, namely, RNA_peptide_98_92_265nm and RNA_pept_U_19min (table 1, herein referred to as method A and B). Both methods use an Aeris 1.7 µm PEPTIDE XB C18 100Å 150x21mm column by Phenomenex. Measurements were performed on a Waters Aqcuity I-Class UPLC.

Method A: RNA_peptide_98_92_265nm

Time	Flowrate (mL/min)	% A	% B	Curve
Initial	0.3	98	2	*Initial*
1	0.3	98	2	6
11	0.3	92	8	6
12.5	0.3	5	95	6
13.5	0.3	5	95	6
15	0.3	98	2	6
16	0.3	98	2	6

Method B: RNA_pept_U_19min

Time	Flowrate (mL/min)	% A	% B	Curve
Initial	0.3	100	0	*Initial*
1	0.3	100	0	6
6	0.3	97	3	6
9	0.3	97	3	6
14	0.3	94	6	6
15.5	0.3	5	95	6
16.5	0.3	5	95	6
18	0.3	100	0	6
19	0.3	100	0	6

Table 1: Two methods used for the UPLC analysis, courtesy of Marcel J. Eleveld, MSc.

2.3 Nuclear Magnetic Resonance Spectroscopy

Nuclear Magnetic Resonance Spectroscopy was performed on a Bruker Ascend 600 MHz spectrometer. The following methods were used: ^{1}H-NMR with water-suppression, ^{13}C-NMR with decoupling, and ^{31}P-NMR with decoupling. All spectra were recorded in D_2O and/or H_2O.

2.4 Mass Spectrometry: Quadrupole Time of Flight

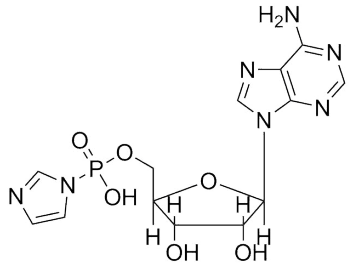
Mass spectrometry was measured on a Xevo G2 QTOF machine, connected to an Acquity H-Class UPLC from Waters. The same column type was used as for the UPLC analysis.

2.5 Monomer activation

The synthesis of phosphorimidazolides with 2,2'-dipyridyldisulphide and triphenylphosphine in DMF was adapted from Lohrmann and Orgel (1978) and Sébastien et al. (2005).[29,30] The same synthesis in DMSO was adapted from Poudyal et al. (2019).[2] The synthesis of phosphoramidates and phosphorimidazolides with trimethylsilyl chloride was adapted from Zhu et al. (2006).[12] Full characterisation of adenosine 5'-monophosphor-(2-methyl)-imidazolide (2-MeImpA), and guanosine 5'-monophosphor-(2-methyl)-imidazolide has been reported by Walton and Szostak (2016).[31] ^{31}P-NMR spectra can be found in **section 7**.

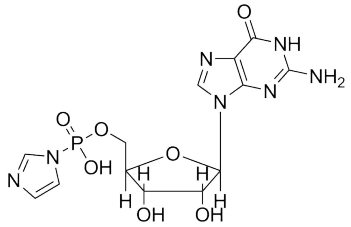
2.5.1 Phosphorimidazolides

Synthesis of ImpA (Adenosine 5'-monophosphorimidazolide)



2,2'-Dipyridyldisulphide (2 eq; 1 mmol; 220 mg), triphenylphosphine (2 eq; 1 mmol; 262 mg), imidazole (5 eq; 2.5 mmol; 170 mg), and triethylamine (13 eq; 6.5 mmol; 0.9 mL) were dissolved in N,N-dimethylformamide (15 mL) in dried glassware under inert atmosphere while stirring vigorously. In a second flask, adenosine 5'-monophosphate monohydrate (1 eq; 0.5 mmol; 174 mg) was suspended in N,N-dimethylformamide (15 mL). Subsequently, the adenosine 5'-monophosphate suspension was added drop-wise over the course of 30 minutes, to the vigorously stirring solution. The reaction mixture was stirred at room temperature for an additional 2.5 hours and monitored by TLC (80$\%_{vol.}$ ethanol in water). The reaction mixture was quenched by rapid addition to a vigorously stirring sodium perchlorate monohydrate (excess; 7.8 mmol; 1.1 g) solution in acetone (100 mL) and diethyl ether (50 mL). The mixture was stirred for 15 minutes and the newly formed suspension was equilibrated over the course of 1 hour. Subsequently, the supernatant was decanted and the crude mixture was purified by means of centrifugation cycles (5000 rpm, 15 min). With each cycle, the supernatant was discarded and replaced with either acetone or diethyl ether followed by vortexing and sonication until a homogeneous suspension was obtained. Thereafter, the product was dried at room temperature under vacuum overnight. The product was isolated as a white powder, and its purity was determined by means of TLC and ^{31}P-NMR. The final product was stored at -70°C as a dried powder in approximately 80% yield. ^{31}P-NMR (240MHz; D_2O) δ: 3.5 (s; minor impurity), -8.0 (s), -11.2 (s; minor impurity) ppm. TLC Rf ~0.4; eluent: 80% ethanol in water; stationary phase: silica gel 60 F_{256} TLC plates.

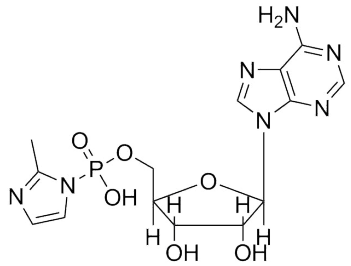
Synthesis of ImpG (Guanosine 5'-monophosphorimidazolide)



Method 1 The same protocol was used as for ImpA. However, adenosine 5'-monophosphate monohydrate was substituted with guanosine 5'-monophosphate monohydrate (1 eq; 0.5 mmol; 181 mg). No conversion of starting material observed.

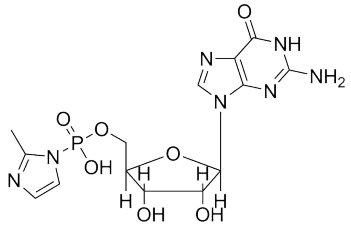
Method 2 To a solution of imidazole (10.5 eq; 2.9 mmol; 197.4 mg) in dimethylsulphoxide (20mL) was added guanosine 5'-monophosphate monohydrate (1 eq; 0.275 mmol; 100mg). The mixture was heated to 50°C until a fine suspension was obtained. Subsequently, the heat was removed and triphenylphosphine (10.5 eq; 2.9 mmol; 750 mg), 2,2'-dipyridyldisulphide (14 eq; 3.9 mmol; 850 mg) and triethylamine (10.5 eq; 2.9 mmol; 1.08 mL) were added to the reaction mixture. The mixture was left stirring for 3 hours after which the reaction was quenched by rapid addition to a vigorously stirring solution of sodium perchlorate (excess; 7.8 mmol; 1.1 g) in acetone (50mL) and diethyl ether (50mL). The mixture was stirred for 15 minutes after which is was left to equilibrate for 1 hour. Thereafter, the supernatant was discarded and the crude powder was purified by centrifugation cycles (5000 rpm, 15 min) and washing with acetone and diethyl ether. The reaction product was analysed by means of ^{31}P-NMR and TLC, demonstrating no conversion of starting material. No ImpG was obtained.

Synthesis of 2-MeImpA (Adenosine 5'-monophosphor-(2-methyl)-imidazolide)



The same protocol was used as for ImpA. However, imidazole was substituted with 2-methyl imidazole (5 eq; 2.5 mmol; 205.25 mg). Product was obtained in approximately 80% yield.^{31}P-NMR (240MHz; D_2O) δ: -7.9 (s), -11.2 (s; minor impurity) ppm. TLC Rf ~0.4; eluent: 80% ethanol in water; stationary phase: silica gel 60 F_{256} TLC plates.

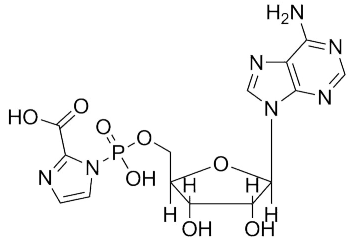
Synthesis of 2-MeImpG (Guanosine 5'-monophosphor-(2-methyl)-imidazolide)



Method 1 The same protocol was used as for ImpA. However, adenosine 5'-monophosphate monohydrate was substituted with guanosine 5'-monophosphate monohydrate (1 eq; 0.5 mmol; 181 mg) and imidazole was substituted with 2-methyl imidazole (5 eq; 2.5 mmol; 205.25 mg). The reaction was unsuccessful, no conversion of starting material was observed.

Method 2 The same protocol was used as for ImpG method 2. However, imidazole was substituted with 2-carboxylic acid imidazole (10.5 eq; 2.9 mmol; 240 mg). The reaction was unsuccessful, no conversion of starting material was observed.

Synthesis of 2-CarboxylImpA (Adenosine 5'-monophosphor-(2-carboxylic acid)-imidazolide

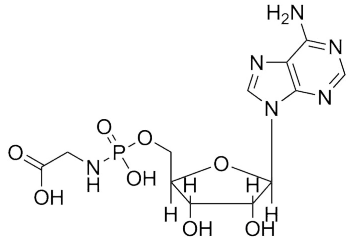


Method 1 The same protocol was used as for ImpA at a smaller scale. However, imidazole was substituted with 2-carboxylic acid imidazole (5 eq; 1.5 mmol; 25.22 mg). The reaction was unsuccessful, no conversion of starting material was observed.

Method 2 To a solution of adenosine 5'-triphosphate disodium salt monohydrate (1 eq; 0.10 mmol; 55.11 mg) in anhydrous pyridine (10mL), in dried glassware under inert atmosphere, was added imidazole-2-carboxylic acid (2 eq; 0.2 mmol; 22.42 mg). The suspension was sonicated for 10 minutes, after which trimethylsilyl chloride (30 eq; 3 mmol; 0.4 mL) was added in a single injection. The suspension turned homogeneously turbid and fumes were present in the flask. The reaction mixture was stirred for 48 hours after which the solvent was evaporated under reduced pressure. The crude mixture was obtained as a yellow viscous oil and analysed by means of ^{31}P-NMR demonstrating large amounts of phosphorus containing contaminants and no product. The crude mixture was not purified further.

2.5.2 Phosphoramidates

Synthesis of adenosine glycine 5'-monophosphoramidate



To a solution of adenosine 5'-triphosphate disodium salt monohydrate (1 eq; 0.4 mmol; 100 mg) in anhydrous pyridine (20mL), in dried glassware under inert atmosphere, was added glycine (2 eq; 0.8 mmol; 30 mg). The suspension was sonicated for 10 minutes, after which trimethylsilyl chloride (27 eq; 10.8 mmol; 1.37 mL) was added in a single injection. The suspension turned homogeneously turbid and fumes were present in the flask. The reaction mixture was stirred for 48 hours after which the solvent was evaporated under reduced pressure. The crude mixture was washed with water and ether, and the aqueous phase was lyophilised overnight. The crude mixture was analysed by means of ^{31}P-NMR and work-up was discontinued since no product was identified.

2.6 In Situ Activation of NMP's

Even though phosphoramidates were not successfully synthesised, another molecule of interest was discovered, namely, poly-L-histidine. It was known in literature that complex coacervates could be formed from poly-L-histidine with some poly-anions[19]. This polymer is of interest since L-histidine has an imidazole moiety as its side-chain. It was argued that combination of poly-L-histine complex coacervates with in situ activation of AMP by these imidazole moieties in the presence of EDC could lead to both partitioning of AMP monomers and provide a form of template for these monomers to react on (figure 4).

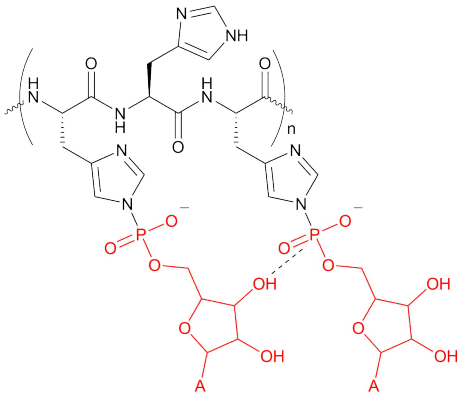


Figure 4: Expected structure of in situ activated AMP with poly-L-histidine, demonstrating the covalent linkages resulting in the proposed templated reaction behaviour between two activated AMP monomers. A = adenine, dashed line represents the close vicinity of functional groups in adjacent activated AMP monomers.

In situ activation of NMP's with imidazole was reported by Burcar et al. in 2015. They reported the use of EDC with imidazole in solution with AMP to spontaneously form short oligomers of AMP and GMP at different pH, and with the potential addition of montmorillonite clay[6]. However, no quantitative assessment of the reaction was performed and only MALDI spectra were reported. With this idea of in situ activation in mind, together with the possibility of poly-L-histidine coacervates experiments were started.

In situ activation of adenosine 5'-monophosphate monohydrate with imidazole analogues
Separate stock solutions of the imidazole analogue of interest, AMP, and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC HCl) were prepared and their pH was adjusted by means of 1000 mM/100 mM NaOH, 1000 mM/100 mM HCl and/or buffers (Tris-Cl 25 mM pH 8; HEPES 20 mM pH 7; MOPS 10 mM pH 6.5; MES 10 mM pH 5; MES 10 mM pH 6). Secondly, 100 µL samples were prepared in the following addition order: 1) Buffer/water and/or salt stock; 2) poly-electrolyte(s) (in case of coacervates); 3) AMP stock; 4) imidazole analogue stock; 5) EDC stock. The vial was vortexed in between additions to prevent concentration gradients. Moreover, concentrations of components differed between experiments. The sample was then monitored by means of UPLC over time.

2.7 Complex coacervation

Complex coacervates were prepared in one of two ways, herein referred to as two component and three component coacervates. The former is the combination of one poly-electrolyte which directly forms complex coacervates with the (modified/activated/non-modified) NMP. The latter refers to the partitioning of (modified/activated/non-modified) NMP in complex coacervates of two poly-electrolytes. Complex Coacervate formation was identified by analysis of centrifuge sediment (precipitate or complex coacervate bulk-phase) and optical microscopy (figure 3). Moreover, partitioning of solutes and oligomerisation of NMP's was analysed by means of UPLC and mass spectrometry.

2.7.1 Two Component Coacervates

Stock solutions of the poly-electrolyte and (modified/unmodified) NMP of interest were prepared separately and adjusted to the right pH by means of 1000 mM/100 mM NaOH, 1000 mM/100 mM HCl and/or buffers (Tris-Cl 25 mM pH 8; HEPES 20 mM pH 7; MOPS 10 mM pH 6.5; MES 10 mM pH 5; MES 10 mM pH 6). Secondly, 100µL samples were prepared by means of the following order of addition: 1) buffer/water and/or salt stock; 2) poly-electrolyte stock; 3) NMP stock. This order was found to aid reproducibility. Vials were vortexed in between additions to prevent concentration gradients. Stock solutions were sonicated for 10min before use.

2.7.2 Three Component Coacervates

Stock solutions of the poly-electrolytes of interest and the NMP of interest were prepared separately and their pH was adjusted by means of 1000 mM/100 mM NaOH, 1000 mM/100 mM HCl and/or buffers (Tris-Cl 25 mM pH 8; HEPES 20 mM pH 7; MOPS 10 mM pH 6.5; MES 10 mM pH 5; MES 10 mM pH 6). Secondly, 100µL samples were prepared by means of the following order of addition: 1) buffer/water stock and/or salt stock; 2) cationic poly-electrolyte stock; 3) anionic poly-electrolyte stock; 4) NMP stock. This order was found to aid reproducibility. Vials were vortexed in between additions to prevent concentration gradients. Stock solutions were sonicated for 10min before use.

2.7.3 Partition Measurements

To analyse the partitioning of NMP's inside complex coacervates two samples are prepared. One containing the complex coacervate mixture, and a control sample without the poly-electrolyte with the same pH, salt concentration, and NMP concentration. Subsequently, both vials are centrifuged at 5000 rpm for 10 minutes after which the supernatant is diluted to the right concentration and injected into the UPLC. Based on the relative peak-area difference between the NMP of interest in the control's supernatant and the complex coacervate sample's supernatant the relative partitioning of NMP's inside the complex coacervate phase was estimated.

3 Results and Discussion

In order to answer the research question laid out in **section 1.3** different types of experiments were performed. These experiments and their results are categorised by three different topics in this section, namely, complex coacervates with NMP's, phosphorimidazolides, phosphoramidates, and poly-L-histidine as the poly-electrolyte. A brief outline of what these sections concern can be found below.

Firstly (**Section 3.1 Complex Coacervates with NMP's**), in order to prepare complex coacervates relevant to this research project, complex coacervates first had to be identified which allowed partitioning of unactivated monomers (NMP's). This section is concerned with the identification of these complex coacervate systems and with the investigation of the partitioning of NMP's into these complex coacervates.

Secondly (**Section 3.2 Phosphorimidazolides**), NMP's without activation are not reactive enough to produce RNA oligomers. Therefore, the compatibility between the identified complex coacervates and the phosphorimidazolide activation method was investigated.

Thirdly (**Section 3.3 Phosphoramidates**), a similar attempt at investigating the compatibility between activation methods and the identified complex coacervates was performed for phosphoramidates.

Lastly (**Section 3.4 Poly-L-Histidine as the Poly-Electrolyte**), as an alternative to pre-activation which is used in the previous two sections in situ activation with poly-L-histidine and NMP's was investigated. Moreover, this is expected to circumvent some issues which came to light in the previous sections.

These categories and their order reflect both the three themes throughout the research project, and to some extent the chronological order in which discoveries were made.

3.1 Complex Coacervates with NMP's

In order to investigate the partitioning of NMP's into the complex coacervate phase, different poly-electrolytes were selected which either formed complex coacervates with each other or formed complex coacervates upon combination with NMP's. As laid out in **section 1.2.1**, complex coacervation is influenced by a large variety of experimental parameters. Therefore, efforts were made to obtain complex coacervates with reasonable size and with as little additives as possible to aid the pre-biotic relevance. Firstly, three component coacervates were investigated.

3.1.1 Three Component Complex Coacervates

Upon mixing complementary poly-electrolytes one can observe three different mixture behaviours, complex coacervation, precipitation, or dissolution (figure 3). Precipitation occurs when interactions between the macro-ions are too large and, therefore, screening is required in order to observe complex coacervation. This can be done by increasing the salt concentration and/or changing the pH. Priftis and Tirrell (2012) found that for combinations of poly-electrolytes which, in the absence of salt, form precipitates an increase of salt concentration resulted in complex coacervation, and after a critical salt concentration the macro-ions were fully dissolved[19]. This methodology was used in combination with pH adjustments, if salt addition had no effect. Based on these preliminary experiments a list of complex coacervate forming mixture compositions was obtained, see table 2.

Macro-anion	Macro-cation	Salt	pH	[Anionic group]	[Cationic group]	[Salt]
PSS	PDADMAC	KBr	7.0	10 mM	10 mM	800 mM
PSS	poly-L-Lysine	KBr	7.0	10 mM	10 mM	800 mM
poly-(D,L)-Glutamic acid	poly-(D,L)-Lysine	none	7.0	6 mM	6 mM	n/a
poly-(D,L)-Glutamic acid	poly-L-Lysine	KBr	7.0	6 mM	6 mM	100 mM
poly-(D,L)-Glutamic acid	oligo-Lysine	KBr	7.0	6 mM	6 mM	100 mM
poly-L-Glutamic acid	poly-(D,L)-Lysine	KBr	7.0	6 mM	6 mM	100 mM
ATP	poly-Allylamine	none	7.0	10 mM	10 mM	n/a
ATP	oligo-Lysine	none	8.0	4 mM	4 mM	n/a
ATP	poly-L-Lysine	none	8.0	4 mM	4 mM	n/a
ATP	poly-(D,L)-Lysine	none	8.0	4 mM	4 mM	n/a
ATP	PDADMAC	none	9.0	5 mM	5 mM	n/a

Table 2: Tabulated mixture compositions for complex coacervation of listed macro-ion combinations. [Anionic group] and [Cationic group] refer to charged moiety concentrations, assuming all moieties are charged.

3.1.2 Two Component Complex Coacervates

After this list of complex coacervate mixtures had been identified the focus of the research was shifted to the identification of poly-electrolytes that were capable of forming complex coacervates with NMP's without the addition of complementary poly-electrolytes. From experiments (table 2) and literature[15,27,28] it was found that nucleic acids (ATP, ADP, AMP and other NMP's) are capable of forming complex coacervates with some poly-electrolytes. Experiments were performed to identify complex coacervate mixtures of single poly-electrolytes with NMP's, the results of which are listed in table 3.

NMP	Macro-cation	Salt	pH	[Anionic group]	[Cationic group]	[Salt]
5'-AMP	poly-Allylamine	none	7.0	5 mM	5 mM	n/a
2'/3'-AMP	poly-Allylamine	none	7.0	5 mM	5 mM	n/a
5'-AMP	poly-L-Lysine	none	8.0	4 mM	4 mM	n/a
5'-AMP	poly-(D,L)-Lysine	none	8.0	4 mM	4 mM	n/a
5'-AMP	poly-L-Histidine	none	5.0	11 mM	25 mM	n/a

Table 3: Tabulated NMP poly-electrolyte mixture compositions exhibiting complex coacervation. Poly-L-histidine as the macro-cation will be discussed in more detail in **section 3.4**. [Anionic group] and [Cationic group] refer to charged moiety concentrations, assuming all moieties are charged.

From these experiments, it was found that tuning system parameters to obtain complex coacervation of adenosine monophosphates with poly-cations was significantly more challenging compared to combinations of two poly-electrolytes. For instance, complex coacervation of poly-lysine with AMP had a narrow pH window of ~7.6 to ~8.8 within which the mixture exhibited complex coacervation. Below and above this pH window, the mixture resulted in precipitation, regardless of the amount of salt present.

3.1.3 Partitioning of NMP's in Complex Coacervates

By combining the data in tables 2 and 3, and based on UPLC partition analysis (see **section 2.7.3**) it was determined that complex coacervates prepared with poly(allylamine), poly(lysine) of different chirality and chain-length, and poly(histidine) are capable of partitioning adenosine monophosphates (figure 5). Since poly-peptides are capable of AMP partitioning, this suggests a naturally occurring link between RNA building blocks and proteins.

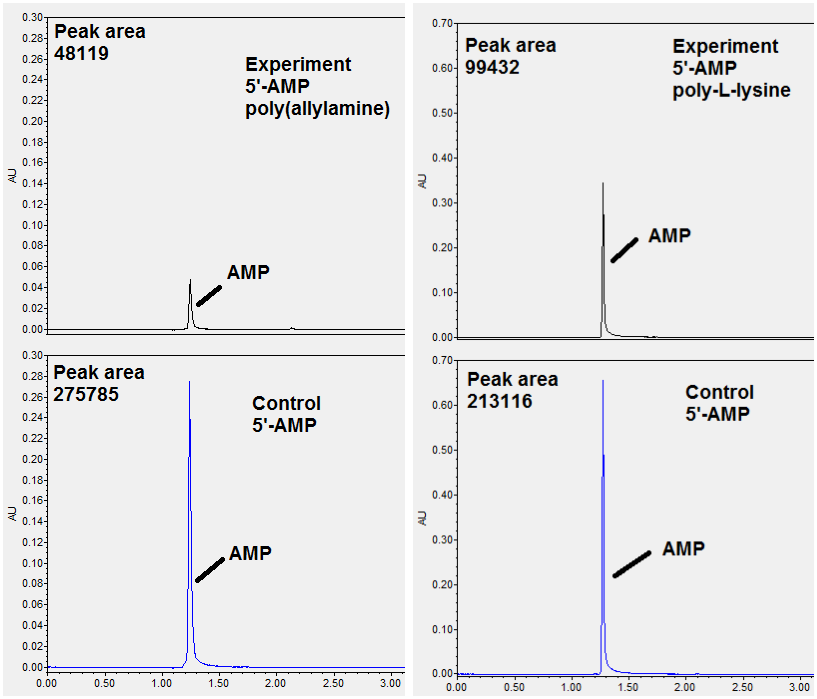


Figure 5: UPLC chromatograms for the analysis of NMP partitioning. Peak area corresponds to the 5'-AMP concentration in the supernatant of the control and experiment samples after centrifugation. Left to right: **1).** For poly(allylamine) (5 mM) and 5'-AMP (5 mM) complex coacervates at pH 7. **2).** For poly-L-lysine (4 mM) and 5'-AMP (4 mM) complex coacervates at pH 7.

At the time of these experiments poly(histidine) had not been identified as a potential candidate for complex coacervation with AMP due to the limited amount of discovered literature on the topic. Therefore, experiments were initially performed on poly(allylamine) and poly(lysine) containing coacervates. Information on poly(histidine) will be presented in **section 3.4**.

Samples of AMP with poly(allylamine) were prepared and analysed over time by means of UPLC. However, due to the relative unreactive nature of unactivated AMP, no reaction products were observed. Therefore, the transition was made to combinations of activated NMP's with poly-cations and potentially poly-anions.

3.2 Phosphorimidazolides

In order to promote oligomerisation of NMP's inside complex coacervates, the monomers had to be chemically activated. Two different methodologies have been used throughout this research project. Phosphorimidazolides, and phosphoramidates. Firstly, phosphorimidazolides were investigated.

3.2.1 Synthesis of Phosphorimidazolides

As laid out in **section 2.5.1** phosphorimidazolides were synthesised by means of a reaction of AMP monohydrate with imidazole in the presence of 2,2'-dipyridyldisulphide and triphenylphosphine in DMF. For adenosine NMP's this reaction resulted in 80% yield with high purity according to TLC (ImpA and 2-MeImpA Rf ~0.4 in $80\%_{vol.}$ EtOH in water) and ^{31}P-NMR (see **section 7**).

2,2'-Dipyridyldisulphide, triphenylphosphine, triethylamine, and the imidazole analogues were all soluble in DMF. However, AMP monohydrate is poorly soluble in the solvent. Therefore, initially, low yields (~30%) were obtained due to the difficulty of transferring the suspension into the reaction mixture in a dropwise fashion. Therefore, to facilitate the addition, use was made of a dropping funnel instead of a syringe. Thereby, clogging was prevented due to the larger opening. The work-up did not result in any problems, and the synthesis can be performed in approximately half a day.

However, when reactions were performed to synthesise ImpG or 2-MeImpG no reaction products were observed. This could be due to the relatively insoluble nature of GMP (less soluble than AMP in DMF and water) in DMF and DMSO. The cause of the failure of these reactions should, therefore, be investigated. Since ImpG and 2-MeImpG could not be synthesised, experiments were carried out solely with ImpA and 2-MeImpA.

3.2.2 Partitioning in Complex Coacervates

The phosphorimidazolides ImpA and 2-MeImpA were used to determine the partitioning of these activated NMP's in complex coacervates. However, it appeared that these activated monomers do not form complex coacervates upon combination with poly-L-lysine and poly(allylamine), or any other poly-cation from table 2. This was argued to arise from the loss of charge-density upon functionalisation of the phosphate group (figure 6).

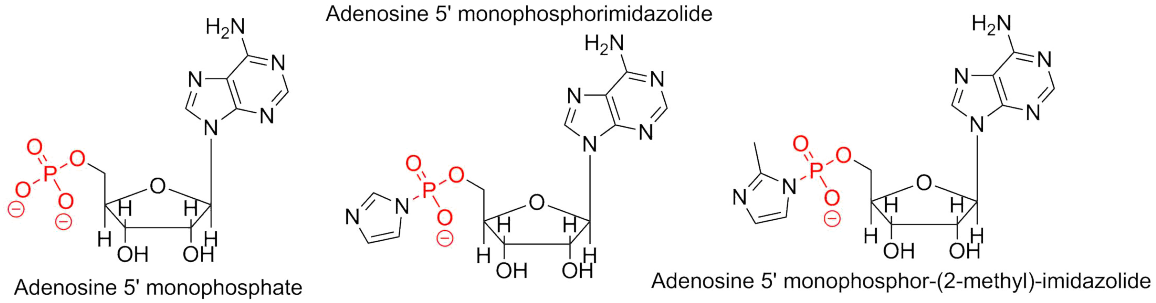


Figure 6: Demonstration of the loss of charge-density, and increase of steric hindrance upon functionalisation of the phosphate group by an imidazole moiety. The phosphate group in red is of interest; upon functionalisation to the phosphorimidazolide one hydroxyl group is substituted and results in the loss of one negative charge.

Moreover, the steric hindrance of the functionalised phosphate group might also play a role in the inability of these compounds to form complex coacervates. The added bulk to the phosphate group could act as a spacer which results in screening of the phosphate charge. Attempts were made to test the role of charge-density on complex coacervation of unactivated AMP. The monomer was combined with PDADMAC (quaternary ammonium compound) and the pH was varied to determine whether it affects the complex coacervate behaviour. However, at no composition and pH were complex coacervates observed. PDADMAC was of interest since this compound's charge-density is independent on pH due to the quaternary ammonium moiety. Therefore, another poly-cation could be used to determine whether charge-density is indeed the driving force behind the lack of complex coacervates for phosphorimidazolides. Moreover, looking at the data from table 2 and 3, each complex coacervate mixture bases its stoichiometry and pH on the major species being divalent AMP, suggesting that complex coacervates do not readily form at lower pH/lower charge-density.

Nonetheless, adenosine phosphorimidazolides do partition to a limited extent into three component complex coacervates with an overall cationic charge. Namely, complex coacervates of PDADMAC and ATP (pH 9, 10 mM and 2.5 mM, respectively) with ImpA (5 mM) (figure 7).

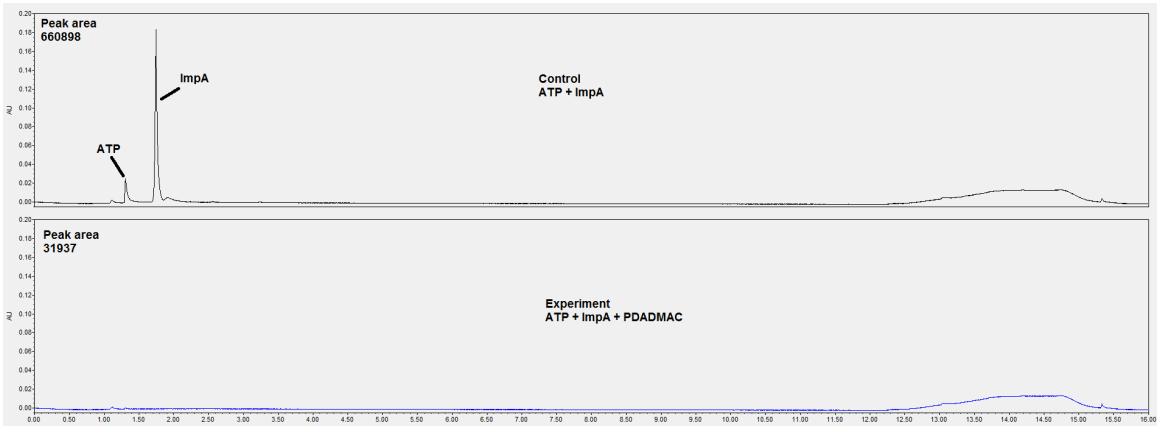


Figure 7: UPLC chromatogram for NMP partitioning analysis. Peak area corresponds to the ImpA and ATP concentration in the control and experiment samples' supernatant after centrifugation for 5 minutes at 5000 rpm. Complex coacervates of PDADMAC (10 mM) and ATP (2.5 mM) with ImpA (5 mM) at pH 9.

Since these materials can partition into complex coacervates, their behaviour inside the complex coacervates was investigated. Experiments were performed to test the oligomerisation behaviour of ImpA. It was found that ImpA formed pyrophosphates in the absence of coacervates, based on mass-spectrometry (see **section 8**). Moreover, upon oligomerisation of NMP's it is expected that both 2'-5' and 3'-5' linkages are formed to some extent. Therefore, multiple peaks should be observable if the product that is being formed is indeed a dimer, instead of a pyrophosphate. However, since no additional peaks were observed it is argued the peak corresponds to pyrophosphates. Moreover, the same behaviour was observed for ImpA and 2-MeImpA when partitioned inside the coacervates. This makes sense considering the mechanism of oligomerisation, which relies on activation through an imidazolium bridged dinucleotide intermediate[32]. In free solution, this is likely not formed competitively with hydrolysis and pyrophosphate products. Therefore, pyrophosphates and the hydrolysis product 5'-AMP are most abundant. For all according spectra and chromatograms see **section 8**.

Based on these experiments it was found that phosphorimidazolides do not oligomerise in detectable yields in free solution, nor do they upon partitioning in coacervates. Moreover, phosphorimidazolides do not form two component complex coacervates due to the lower charge-density compared to 5'-AMP. However, they can be partitioned into three-component complex coacervates. Furthermore, based on UPLC analysis it appears that ImpA hydrolyses faster than 2-MeImpA because of the rapid depletion of the ImpA peak compared to 2-MeImpA (**section 8**). In order to circumvent the charge-density issue and to use another means of activation, phosphoramidates were investigated.

3.3 Phosphoramidates

An alternative activation method for NMP's is the use of phosphoramidates. These activating groups rely on the same P-N functionalisation as phosphorimidazolides. Moreover, an additional negative charge can be part of their scaffold theoretically negating the charge-density effect. However, they lack the aromatic scaffold which is necessary for the high effectiveness of phosphorimidazolides for primer extension[32]. Nonetheless, attempts were made to prepare phosphoramidates of glycine and AMP. Unfortunately, no product was isolated (^{31}P-NMR spectra in **section 7**). This can partly be attributed to the limited solubility of the compounds in pyridine, the water used in the washing step leading to possible hydrolysis, and the rupture of the flask in the freeze-dryer. Overall, this synthesis was not successful and the work-flow leaves room for significant improvements.

Since the investigation of this activation method had thus far not been fruitful the shift was made to in-situ activation of NMP's with poly-L-histidine.

3.4 Poly-L-Histidine as the Poly-Electrolyte

3.4.1 Three Component Complex Coacervates

Firstly, it was researched whether poly-L-histidine could form complex coacervates with poly-(D,L)-glutamic acid. (D,L)-glutamic acid was used since it is known from literature that two enantiopure poly-peptides do not form complex coacervates[17]. A number of complex coacervate systems was identified for poly-L-histidine, table 4. However, during preparation of these complex coacervates, it became apparent that adjusting the pH after mixing resulted in precipitates in all cases. Whereas, when buffered stock solutions were combined complex coacervates would form. This is believed to be due to the high stability of precipitates, and the inability of protons to penetrate the solid precipitate effectively.

Macro-anion	Macro-cation	Salt	pH	[Anionic group]	[Cationic group]	[Salt]
PSS	p-L-Histidine	KBr	5	5 mM	5 mM	>800 mM
p-(D,L)-Glutamic acid	p-L-Histidine	none	5	5 mM	5 mM	n/a
p-(D,L)-Aspartic acid*	p-L-Histidine*	none	5	5 mM	5 mM	n/a

Table 4: Mixture compositions of poly-L-histidine - macro-anion complex coacervates. *reported in literature by Tan et al. (2013)[33]. Samples were prepared in pH 5 MES buffer 10 mM stock solution and pH was adjusted with NaOH and HCl 100 mM solutions. pH of stock solutions was adjusted to 5 before mixing.

3.4.2 Three Component In Situ Activation

Experiments were performed with poly-L-histidine (5 mM), poly-(D,L)-glutamic acid (2.5 mM) and AMP (5 mM) in the presence of EDC (100 mM) at pH 5 in 10 mM MES buffer. However, no oligomerisation was observed. Moreover, it appeared that AMP was not fully dissolved in the mixture and that the pH was not constant during measurements. Therefore, new experiments were performed. All stock solutions were adjusted to the right pH with buffer and hydrochloric acid (100 mM stock), sonicated before use, and combined in the required compositions. However, due

to the addition order, it was found that when the AMP and poly-L-histidine stock were adjusted to pH 5 and sonicated for 10 minutes before use, two component complex coacervates spontaneously formed upon mixing (figure 8), as mentioned in table 3.

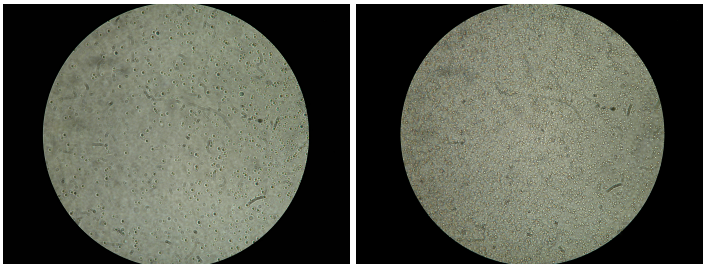


Figure 8: Optical micro-graphs of two component complex coacervates of poly-L-histidine (25 mM) and AMP (11 mM) at pH 5 with 10 mM MES stock. (Circle is 1.5mm in diameter). Figures are of different parts of the droplet under the optical microscope demonstrating heterogeneity in complex coacervate size.

Therefore, it was decided to omit poly-(D,L)-glutamic acid from the experiment and focus on the two component complex coacervate system of poly-L-histidine and AMP in the presence of EDC.

3.4.3 Two Component In Situ Activation

Two component complex coacervates of poly-L-histidine (15 mM) and 5'-AMP (11 mM) were prepared and EDC (100 mM) was added to the solution. Subsequently, the mixtures were vortexed and left stationary at room temperature over time. UPLC analysis was performed after 5, and 16 days to determine how the samples changed over time. Moreover, the first sample was measured using UPLC method A and the second measurement using method B (table 1), since this gave better separation. The samples were prepared according to table 5. The samples were prepared from the following stock solutions: 5'-AMP (20 mM, pH 5), poly-L-histidine (100 mM monomer, pH 5), EDC (2000 mM, pH 5), and MES buffer (10 mM, pH 5). Moreover, all stock solutions were sonicated before use.

Sample	*[5'-AMP]*	*[poly-L-histidine monomer]*	*[EDC]*	*Total volume (µL)*
A	11 mM	25 mM	100 mM	100 µL
B	0	25 mM	100 mM	100 µL
C	11 mM	0	100 mM	100 µL
D	11 mM	25 mM	100 mM	100 µL
E	11 mM	25 mM	100 mM	100 µL

Table 5: Sample preparation for the analysis of in-situ activation of 5'-AMP in the presence of poly-L-histidine and EDC. Complex coacervates of AMP and poly-L-histidine are formed at these compositions. [compound] correspond to concentration and imidazole moiety concentration in the case of poly-L-histidine.

As can be seen from figure 9 a limited amount of pyrophosphate is formed, which is confirmed by mass spectrometry. Moreover, some smaller humps are visible on the chromatogram. However, these peaks were not observable in the QTOF. Therefore, it is argued that, although it seems that the AMP is being activated, it results in the formation of a limited amount of pyrophosphates.

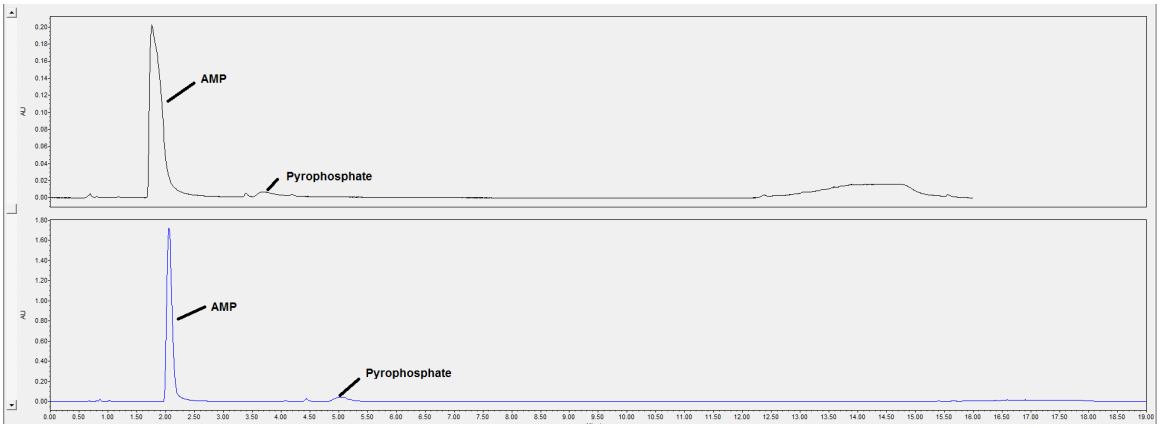


Figure 9: UPLC Chromatogram of two component complex coacervate mixture of 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100 mM) at pH 5 (sample A). Measured 5 and 16 days after start, respectively.

There are two possible explanations for these observations. It could be that the activated AMP readily hydrolysis before any other product is observed, and/or that the reaction is sufficiently slow such that only a limited amount of monomers are activated at any time. Moreover, the in situ activation does not seem to compete with pre-activated phosphorimidazolides, which demonstrated higher yields of pyrophosphates (**section 8**).

However, during data analysis it appeared that one sample exhibited anomalous behaviour compared to its duplicates, namely, sample E (figure 10). As seen in table 5 this sample was of the same composition as sample A and D. However, it appeared that after 16 days a different peak emerged with an m/z value of 312. A mass which is lighter than monomer. Moreover, from the UV-spectra from the UPLC, it appeared that it has the same absorption spectrum as adenine. Therefore, it is argued that the compound should contain adenine.

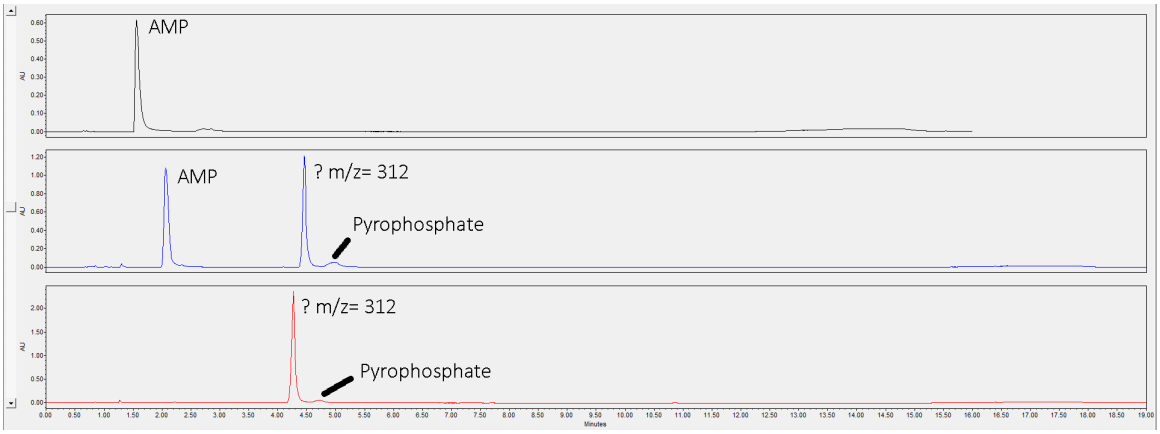


Figure 10: UPLC Chromatogram of two component complex coacervate mixture of 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100 mM) at pH 5 (sample E). Measured 5, 16 and 21 days after start, respectively.

In order to determine what the compound could be ^{1}H-NMR and ^{31}P-NMR spectra were recorded of samples A and E (**section 7**). From the ^{31}P-NMR spectra no significant observable difference could be identified. However, the ^{1}H-NMR spectrum demonstrated the absence of a peak at δ 7.05 ppm in sample E. This peak corresponds to the primary amine on adenine. Therefore, based on mass and on ^{1}H-NMR spectroscopy the unknown product is expected to be an adduct of adenine and MES buffer (figure 11). Even though significant amounts of this side product have been observed in sample E, neither A nor D demonstrated the same behaviour. It is, therefore, argued that this formation could be due to experimental error. However, this still raises the question of how the product could have been formed.

3.4.4 Adduct Formation

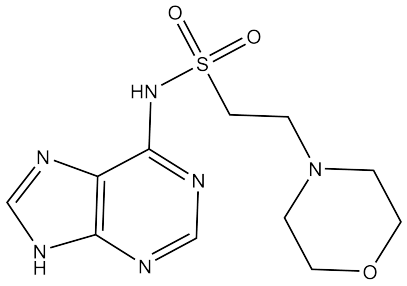


Figure 11: Proposed structure based on mass and ^{1}H-NMR spectroscopy of the compound corresponding to the unknown peak in the UPLC chromatogram of sample E (figure 10). It is argued the compound is an adduct of adenine with MES buffer.

It has been reported by Gates (2006)[34] that deglycosylation of AMP can occur relatively quickly ($t_{1/2}$ of 3h) upon alkylation of an aromatic nitrogen atom in adenosine[34]. However, as seen in figure 11, it is the primary amine which is reacted with the sulphonate group on the MES buffer towards a sulphonamide. The nitrogen in this functional group is electron withdrawing due to a resonance structure in which the nitrogen bears a positive charge[35]. Moreover, since this nitrogen atom is attached to the aromatic scaffold of adenine, it acts as an inductive electron withdrawing group which results in an electron deficient aromatic scaffold[35]. Moreover, protonation of one or more of the adenine's aromatic nitrogen atoms will result in more electron withdrawing behaviour, which could resemble the effect of aromatic nitrogen alkylation. With these effects in mind, it is proposed that the MES:adenine adduct is formed by means of the following mechanism; Firstly, the primary amine performs a nucleophilic attack on the sulphonate group after which water acts as a leaving group resulting in the sulphonamide. Secondly, the added inductive withdrawing effect in combination with protonation of aromatic nitrogen atoms results in an electron deficient aromatic scaffold which can be alleviated by means of deglycosylation forming the MES:adenine adduct. Thirdly, the oxonium ion is hydrolysed and results in a mixture of pentose sugars of different conformations with a 5'-phosphate group. (figure 12).

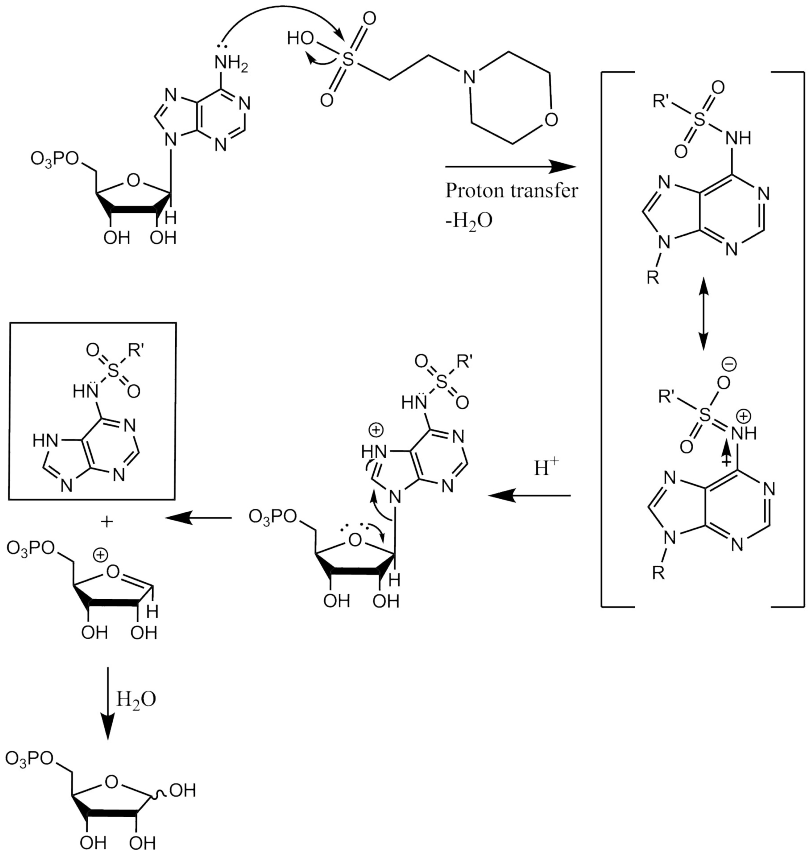


Figure 12: Proposed mechanism for the MES:adenine adduct formation.

Even though no oligomerisation was observed, poly-L-histidine is still a candidate due to the low concentrations that were used. If higher concentrations were used for AMP compared to poly-L-histidine it is argued to prevent unoccupied imidazole moieties aiding the template effect the poly-L-histidine could have on oligomerisation. Moreover, this would make analysis by means of NMR more feasible. Furthermore, the ex-situ activation of L-histidine and poly-L-histidine with AMP should be attempted to determine the feasibility of this activation method. Lastly, the formation of an MES:adenine adduct can be prevented by choosing a different buffer. The NMR analysis can be found in **section 7** and the UPLC and Mass analysis can be found in **section 9**.

4 Conclusion and Outlook

We investigated whether complex coacervates can act as pre-biotic compartments for RNA oligomerisation. Even though it has been demonstrated that complex coacervates can act as a means of compartmentalisation for RNA monomers, whether these be activated or canonical, oligomerisation of these monomers has not been observed.

We demonstrated that three and two component complex coacervates can be formed with canonical 2',3' and 5'-AMP and a variety of poly-cations. However, it appeared that the activated phosphorimidazolides were not capable of forming two component complex coacervates, due to charge density issues. Phosphorimidazolides of AMP could be synthesised with high yield and purity according to ^{31}P-NMR and TLC. However, phosphorimidazolides of GMP and of 2-carboxyl imidazole were not synthesised successfully. The former can be attributed to poor solubility, and the latter to the added negative charge which could cause side-reactions. Furthermore, synthesis of the glycine adenosine phosphoramidate yielded no product, either due to rapidly onset hydrolysis, or inadequate protocol. Additionally, the in situ activation of AMP with poly-L-histidine and EDC yielded limited amounts of pyrophosphates and in an anomalous case the production of a major side-product expected to be an adduct of adenine and MES buffer.

As an outlook, we argue that the in situ activation of imidazole with poly-L-histidine is still a promising strategy to obtain RNA oligomerisation products, regardless of the adduct formation. This problem can easily be circumvented by using a different buffer. Furthermore, the experiment should be performed at higher concentrations to allow for more effective analysis by means of NMR spectroscopy, and to make sure that each imidazole moiety can be activated to prevent unoccupied gaps. Moreover, the poly-L-histidine could be used to pre-activate the AMP's with 2,2'-dipyridyldisulphide and triphenylphosphine in an ex-situ reaction to determine whether this reaction is likely to occur at all, and to see whether the products oligomerise upon solvation in water. Furthermore, different activation methodologies than phosphorimidazolides should be investigated. For instance, OAt esters, and cyclic phosphates should be tested more rigorously.

Overall, the presented combinations of complex coacervates and activation methods do not suggest that pre-biotic oligomerisation of RNA is feasible in complex coacervates. However, since this thesis mainly covers one of many different RNA activation methods more research is necessary to fully determine whether the phenomenon could play a role in pre-biotic compartmentalisation for RNA oligomerisation.

5 Acknowledgements

I would like to thank, in particular:

Marcel J. Eleveld, MSc for the daily supervision, helpful discussions, and feedback on this thesis.
Prof. dr. Sijbren Otto for allowing me to perform the research project in his group and for feedback on this thesis.

Furthermore, I would like to thank the members of the Otto-group for the wonderful atmosphere and the help I have received during my research project.

6 References

[1] Strulson, C. A.; Molden, R. C.; Keating, C. D.; Bevilacqua, P. C. *Nat. Chem.* **2012**, *4*, 941–946.
[2] Poudyal, R. R.; Guth-Metzler, R. M.; Veenis, A. J.; Frankel, E. A.; Keating, C. D.; Bevilacqua, P. C. *Nat. Comm.* **2019**, *10*, 1–13.
[3] Noller, H.; Hoffarth, V.; Zimniak, L. *Science* **1992**, *256*, 1416–1419.
[4] Oró, J.; Stephen-Sherwood, E. *Origins of Life* **1974**, *5*, 159–172.
[5] Ferris, J. P.; Ertem, G. *Science* **1992**, *257*, 1387–1389.
[6] Burcar, B. T.; Jawed, M.; Shah, H.; McGown, L. B. *Orig. Life Evol. Biosph.* **2015**, *45*, 31–40.
[7] Joyce, G. F. *Nature* **2002**, *418*, 214–221.
[8] Renz, M.; Lohrmann, R.; Orgel, L. E. *BBA - Nucleic Acids and Protein Synthesis* **1971**, *240*, 463–471.
[9] Verlander, M. S.; Lohrmann, R.; Orgel, L. E. *J. Mol. Evol.* **1973**, *2*, 303–316.
[10] Šponer, J. E.; Šponer, J.; Di Mauro, E. *J. Mol. Evol.* **2016**, *82*, 5–10.
[11] Hagenbuch, P.; Kervio, E.; Hochgesand, A.; Plutowski, U.; Richert, C. *Angew. Chem. Int. Ed.* **2005**, *44*, 6588–6592.
[12] Zhu, J.; Fu, H.; Jiang, Y.; Zhao, Y. *J. Org. Chem.* **2006**, *71*, 1722–1724.
[13] Prabahar, K. J.; Ferris, J. P. *J. Am. Chem. Soc.* **1997**, *119*, 4330–4337.
[14] Bungenberg de Jong, H. G.; Kruyt, H. R. *Proc. Kon. Ned. Akad. Wetensch.* **1929**, *32*, 849–856.
[15] Frankel, E. A.; Bevilacqua, P. C.; Keating, C. D. *Langmuir* **2016**, *32*, 2041–2049.
[16] Tang, T. Y. D.; Antognozzi, M.; Vicary, J. A.; Perriman, A. W.; Mann, S. *Soft Matter* **2013**, *9*, 7647–7656.
[17] Perry, S. L.; Leon, L.; Hoffmann, K. Q.; Kade, M. J.; Priftis, D.; Black, K. A.; Wong, D.; Klein, R. A.; Pierce, C. F.; Margossian, K. O.; Whitmer, J. K.; Qin, J.; de Pablo, J. J.; Tirrell, M. *Nat. Comm.* **2015**, *6*, 6052.
[18] Spruijt, E. *Wageningen University Thesis* **2012**,
[19] Priftis, D.; Tirrell, M. *Soft Matter* **2012**, *8*, 9396–9405.
[20] Starchenko, V.; Müller, M.; Lebovka, N. *J. Phys. Chem. B B* **2012**, *116*, 14961–14967.
[21] Kayitmazer, A. B. *Adv. Colloid Interface Sci.* **2017**, *239*, 169–177.
[22] Nichols, M. K. *University of Bristol department of Chemistry, PhD thesis* **2018**,
[23] Yang, B.; Jin, S.; Park, Y.; Jung, Y. M.; Cha, H. J. *Small* **2018**, *14*, 1–12.
[24] Abel, J. S.; Stangle, G. C.; Schilling, C. H.; Aksay, I. A. *J. Mater. Res.* **1994**, *9*, 451–461.
[25] Priftis, D.; Tirrell, M. *Soft Matter* **2012**, *8*, 9396–9405.
[26] Drobot, B.; Iglesias-Artola, J. M.; Le Vay, K.; Mayr, V.; Kar, M.; Kreysing, M.; Mutschler, H.; Tang, T.-y. Y. *Nat. Comm.* **2018**, *9*, 1–9.
[27] Jia, T. Z.; Hentrich, C.; Szostak, J. W. *Orig. Life Evol. Biosph.* **2014**, *44*, 1–12.
[28] Koga, S.; Williams, D. S.; Perriman, A. W.; Mann, S. *Nat. Chem.* **2011**, *3*, 720–4.
[29] Lohrmann, R.; Orgel, L. E. *Tetrahedron* **1978**, *34*, 853–855.
[30] Sébastien, P.; Mariana, L.-Q.; Thomas, T. *Curr. Protoc. Mol. Biol.* **2005**, *72*, 1–26.
[31] Walton, T.; Szostak, J. W. *J. Am. Chem. Soc.* **2016**, *138*, 11996–12002.
[32] Walton, T.; Zhang, W.; Li, L.; Tam, C. P.; Szostak, J. W. *Angew. Chem. Int. Ed.* **2019**,
[33] Tan, Y.; Yildiz, U. H.; Wei, W.; Waite, J. H.; Miserez, A. *Biomacromolecules* **2013**, *14*, 1715–1726.
[34] Gates, K. S. *Reviews of Reactive Intermediate Chemistry* **2006**, 333–378.
[35] Chataigner, I.; Panel, C.; Gérard, H.; Piettre, S. R. *Chem. Comm.s* **2007**, 3288.

7 Appendix A

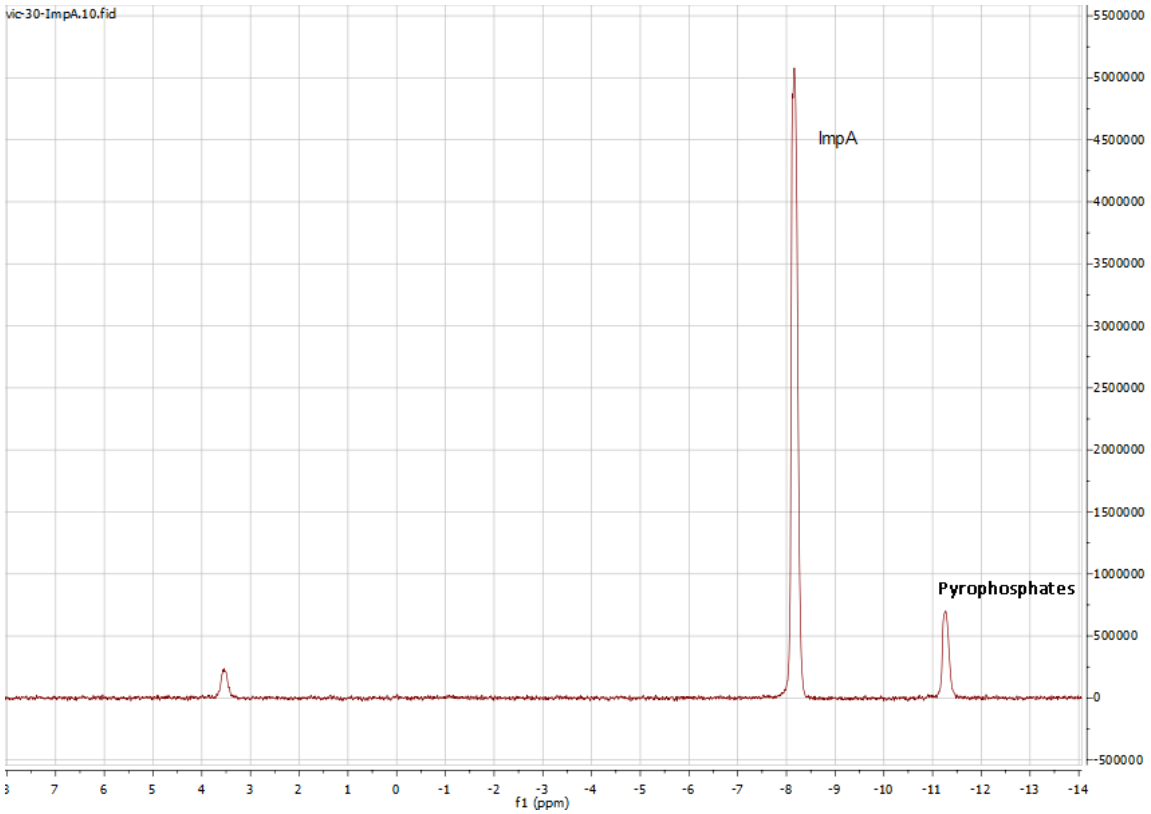


Figure 13: Phosphorus NMR spectrum of ImpA, pyrophosphate peak is visible due to use of water as the solvent.

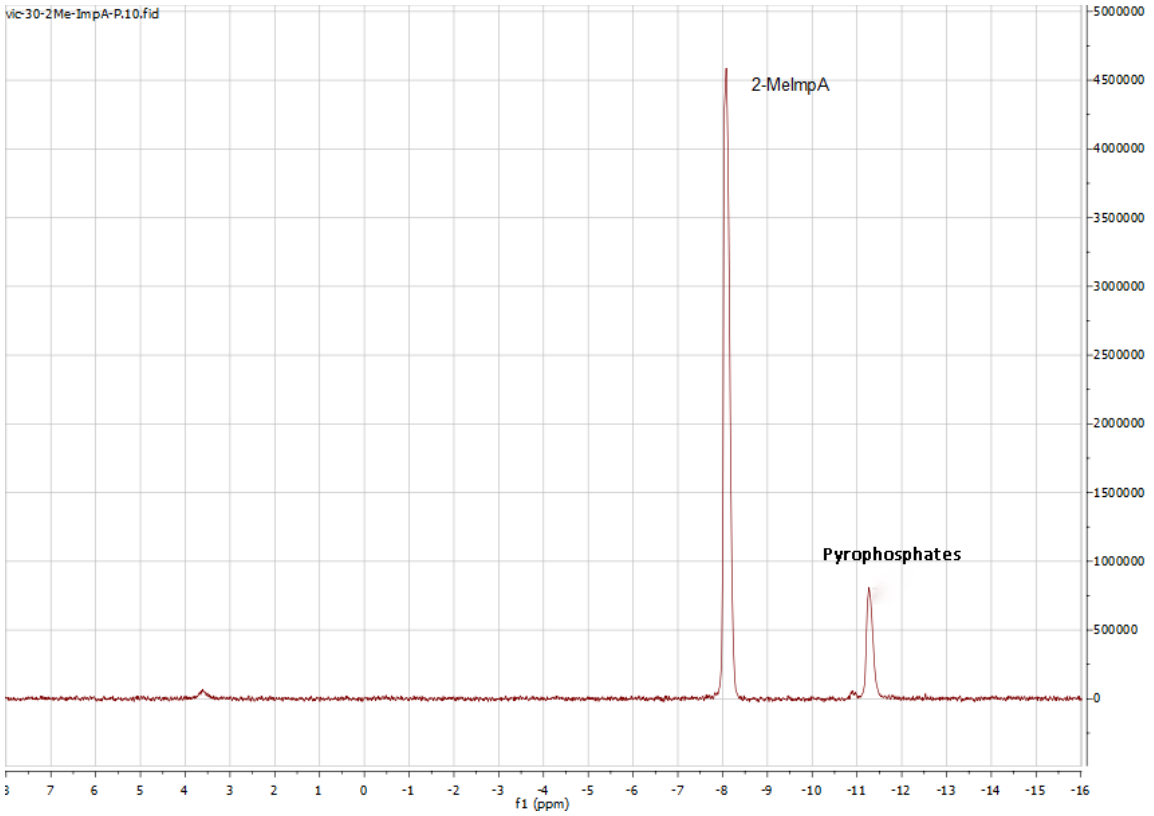


Figure 14: Phosphorus NMR spectrum of 2-MeImpA, pyrophosphate peak is visible due to use of water as the solvent.

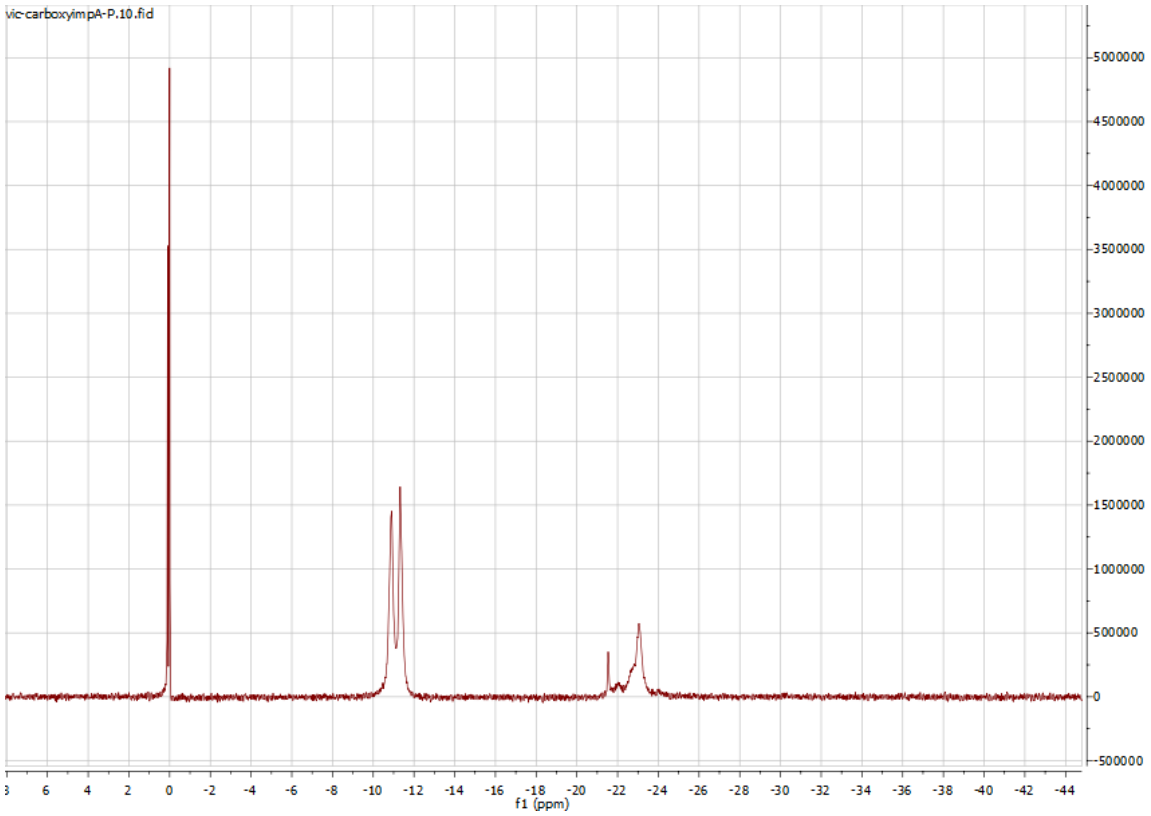


Figure 15: Phosphorus NMR spectrum of 2-carboxyl ImpA demonstrating that the crude product is contaminated with other phosphorous containing material.

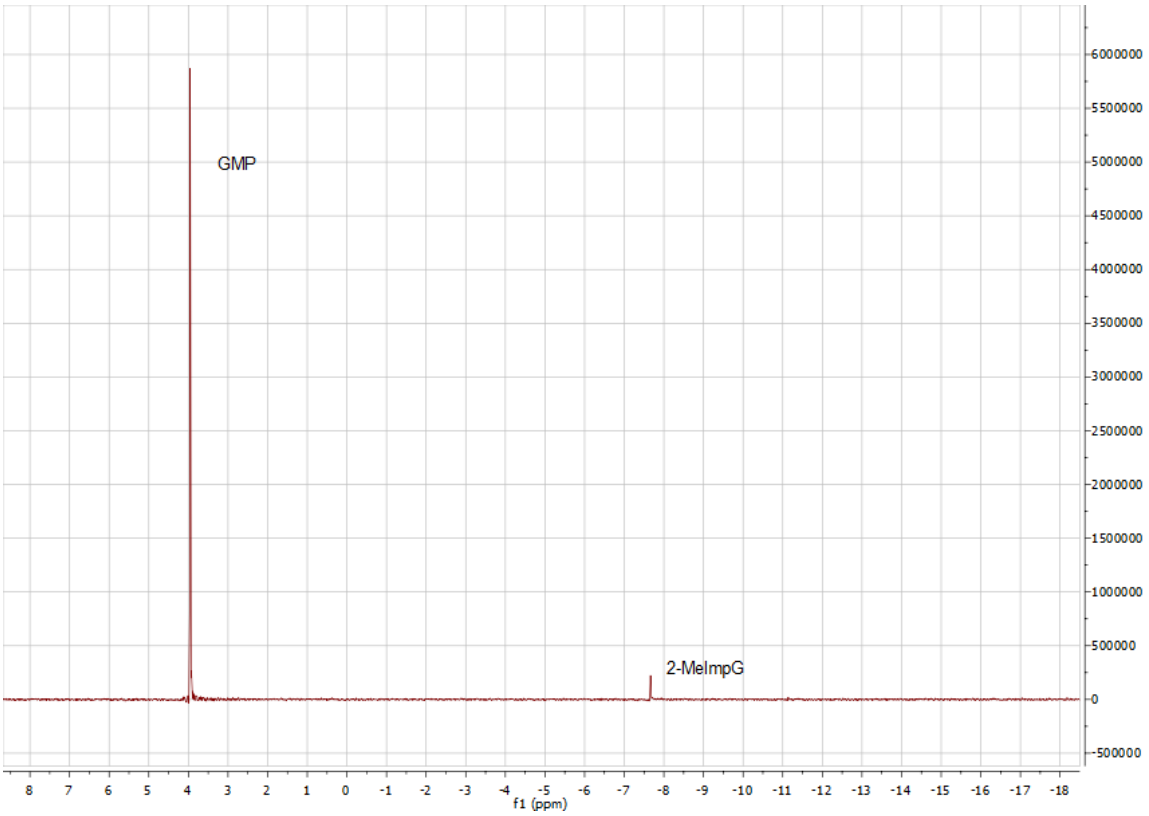


Figure 16: Phosphorus NMR spectrum of 2-MeImpG demonstrating the low starting material conversion.

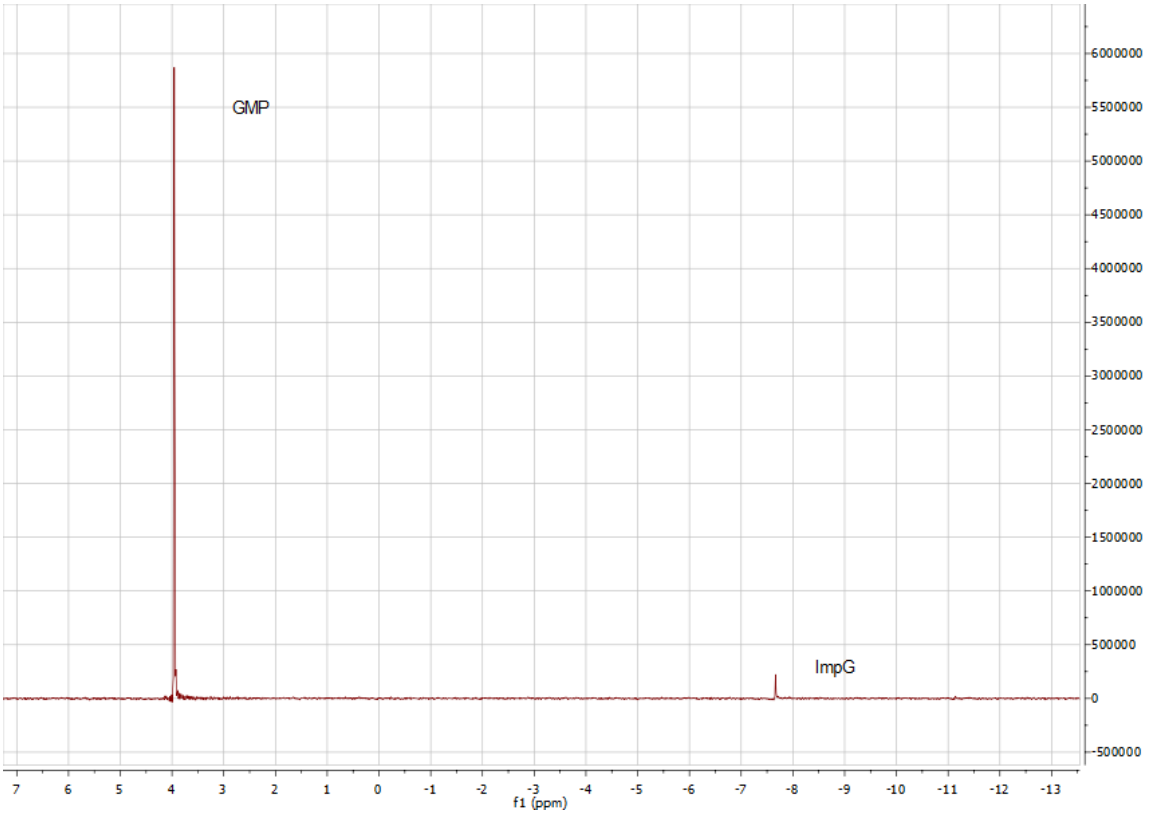


Figure 17: Phosphorus NMR spectrum of ImpG demonstrating the low starting material conversion.

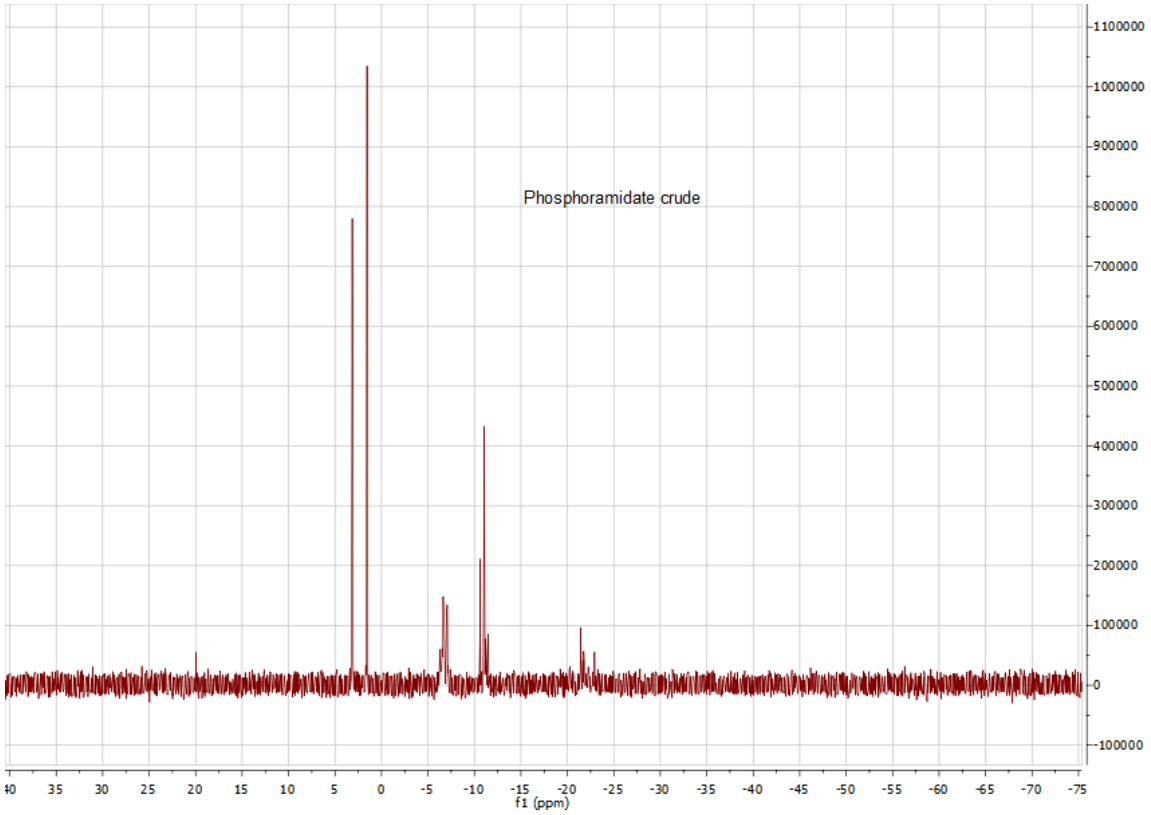


Figure 18: Phosphorus NMR spectrum of adenosine glycine phosphoramidate demonstrating the large amount of phosphorus containing contaminants.

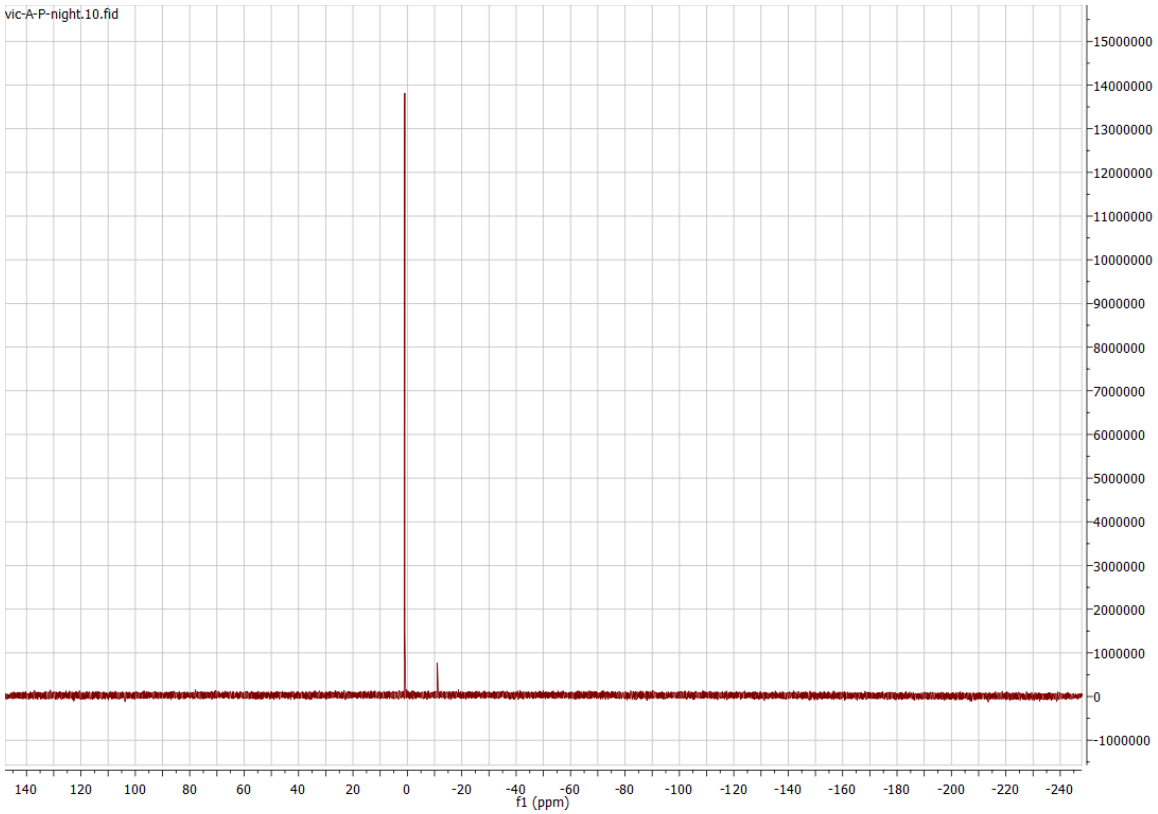


Figure 19: Phosphorus NMR spectrum of in situ activated two component complex coacervate mixture (sample A) containing 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

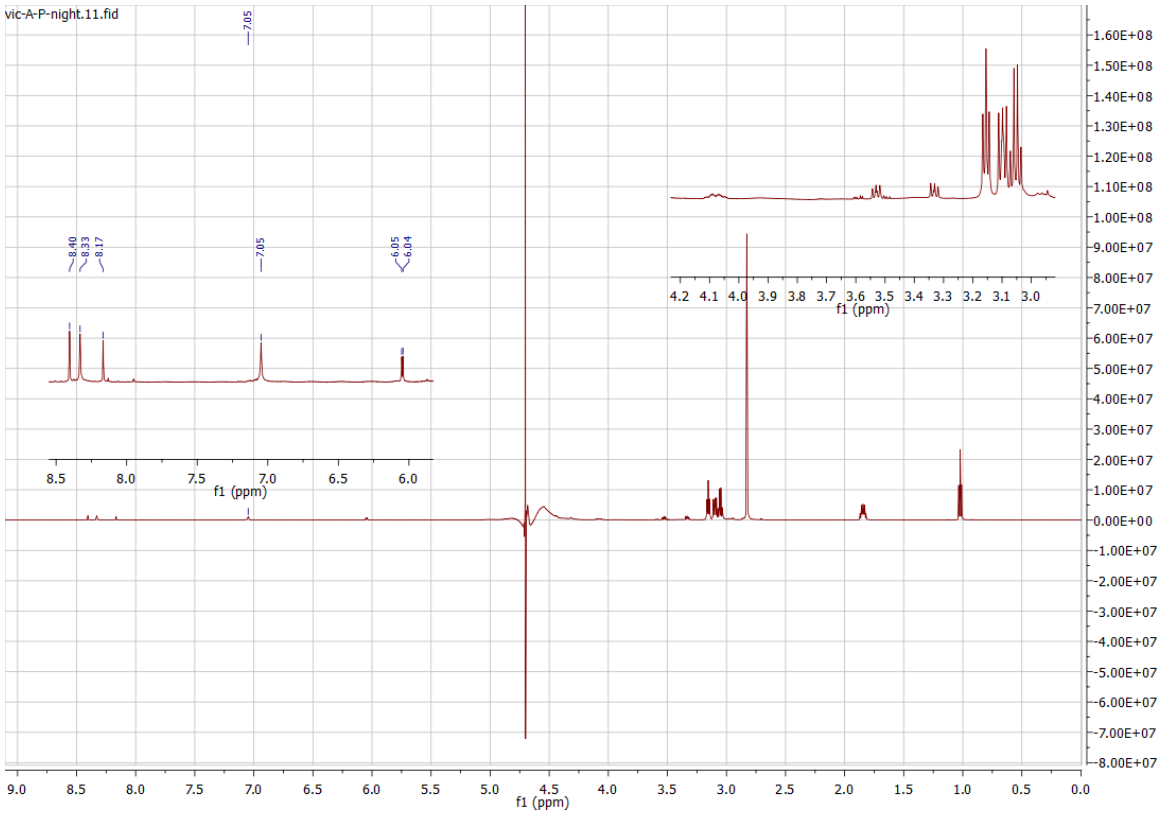


Figure 20: Proton NMR spectrum of in situ activated two component complex coacervate mixture (sample A) containing 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

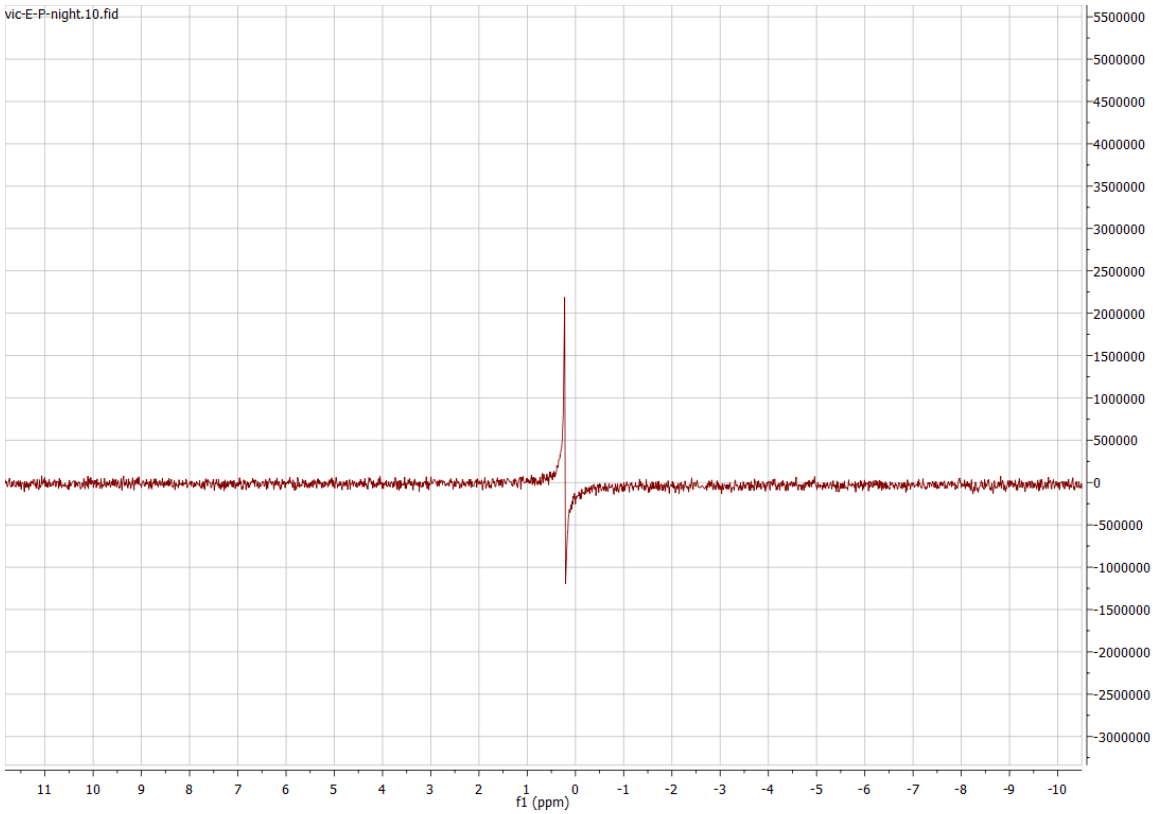


Figure 21: Phosphorus NMR spectrum of in situ activated two component complex coacervate mixture (sample E) containing 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

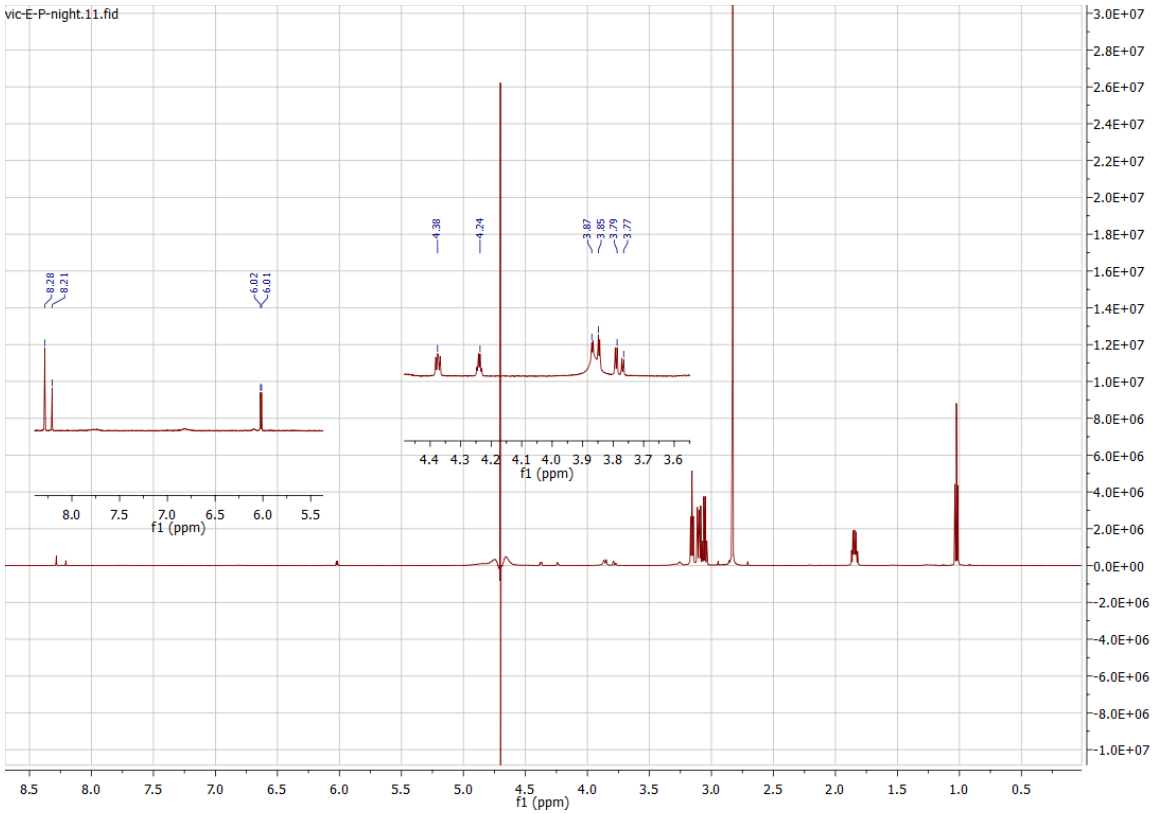


Figure 22: Proton NMR spectrum of in situ activated two component complex coacervate mixture (sample E) containing 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

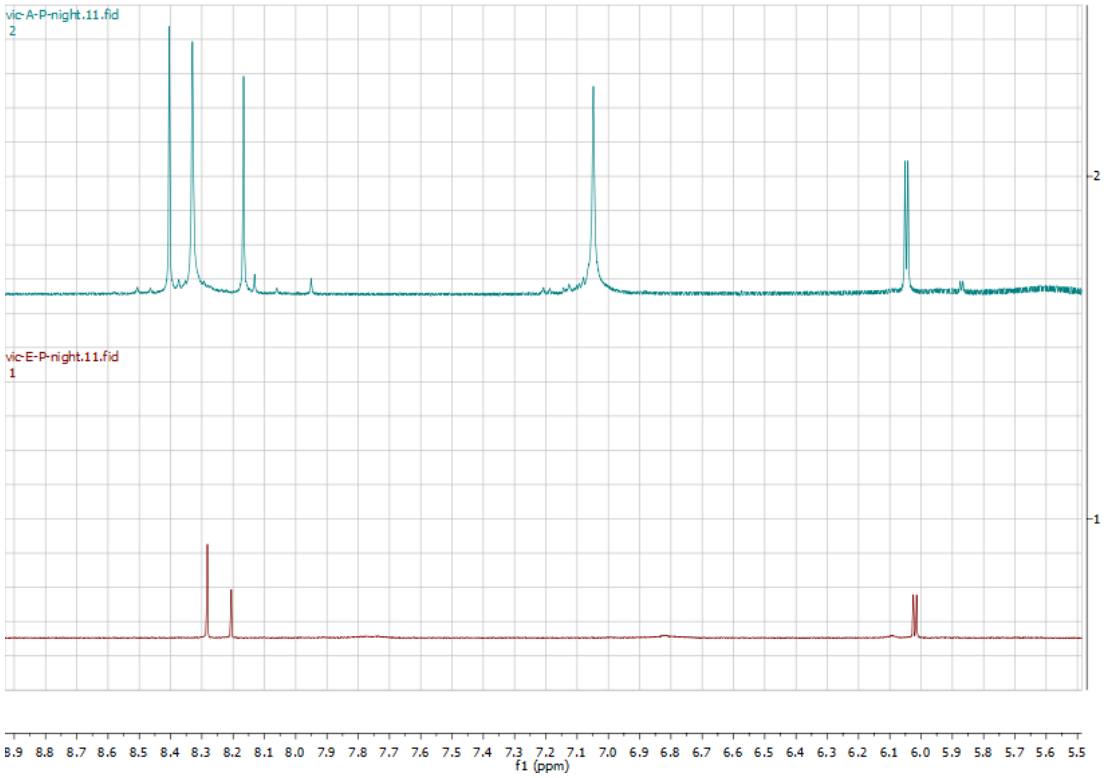


Figure 23: Comparison proton NMR segment of sample A and sample E, demonstrating the loss of a singlet at δ 7 ppm. Both samples contain 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

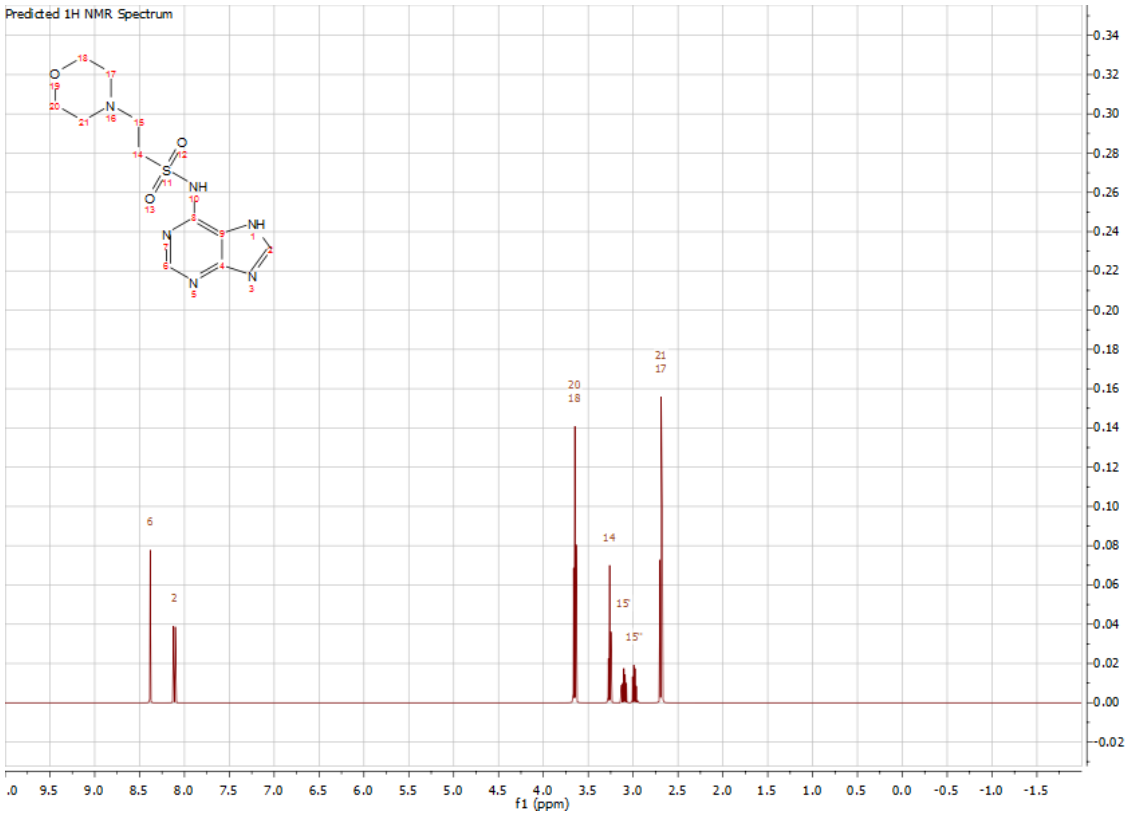


Figure 24: Predicted NMR spectrum of the MES:adenine adduct.

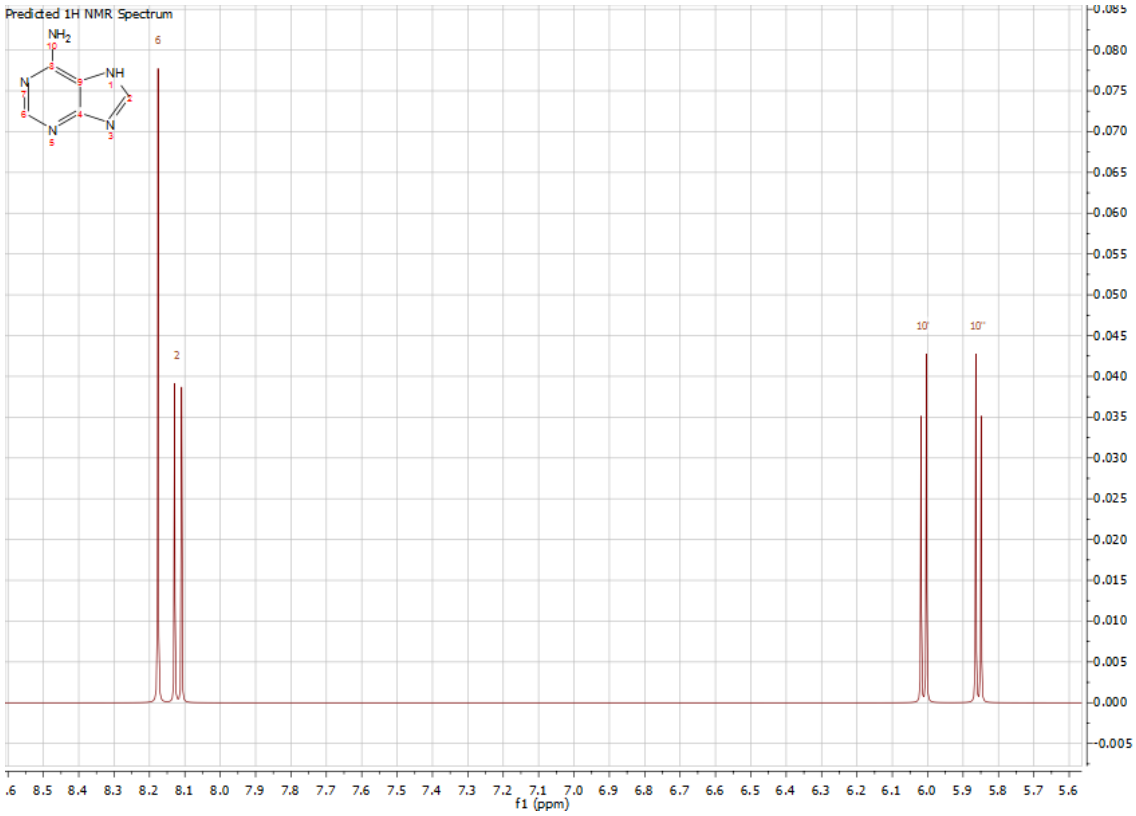


Figure 25: Predicted NMR spectrum of adenine.

8 Appendix B

The first two UPLC measurements per experiment were performed using UPLC method A (RNA_peptide_98_92_265nm) and the latter two using UPLC method B (RNA_pept_U_19min) since this method gave better separation (table 1). Peak area was not measured for these samples since the same trend was observable for the samples, namely, hydrolysis and pyrophosphate formation as major processes.

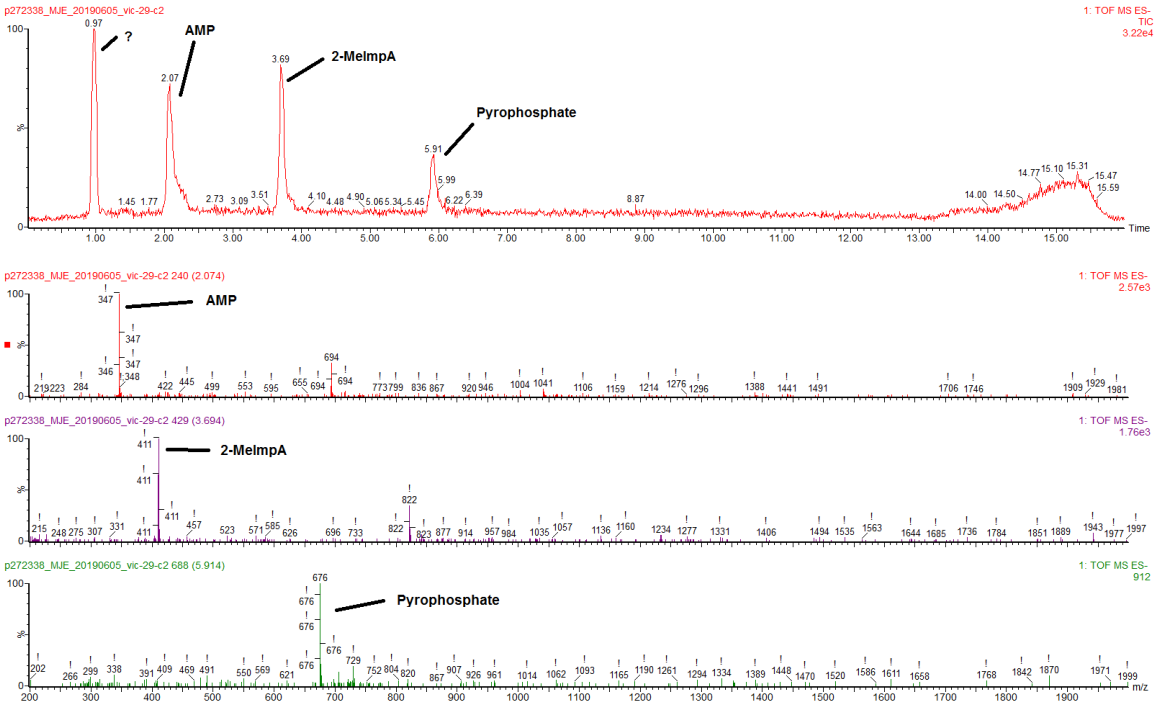


Figure 26: Mass spectrometry analysis to identify peaks at different retention times for the UPLC analysis of the samples below. Sample initially contained PDADMAC (10 mM) and ATP (2.5 mM) with ImpA (5 mM) at pH 9, measured after 21 days.

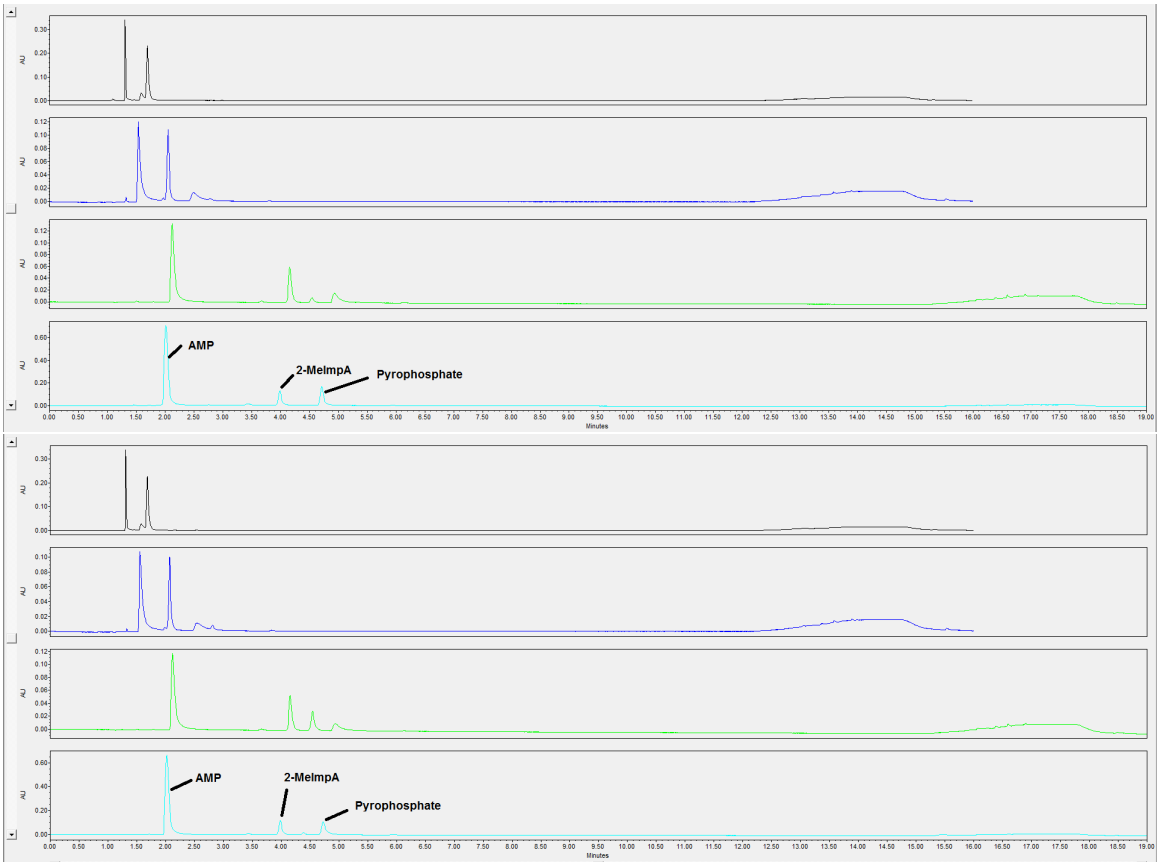


Figure 27: Sample C1 over time. Top four chromatograms 2.5 mM ATP + 5 mM 2-MeImpA. Bottom chromatogram 10 mM PDADMAC + 2.5 mM ATP + 5 mM 2-MeImpA. First measurement after 4 days, second after 10 days, third after 12 days, and fourth after 21 days.

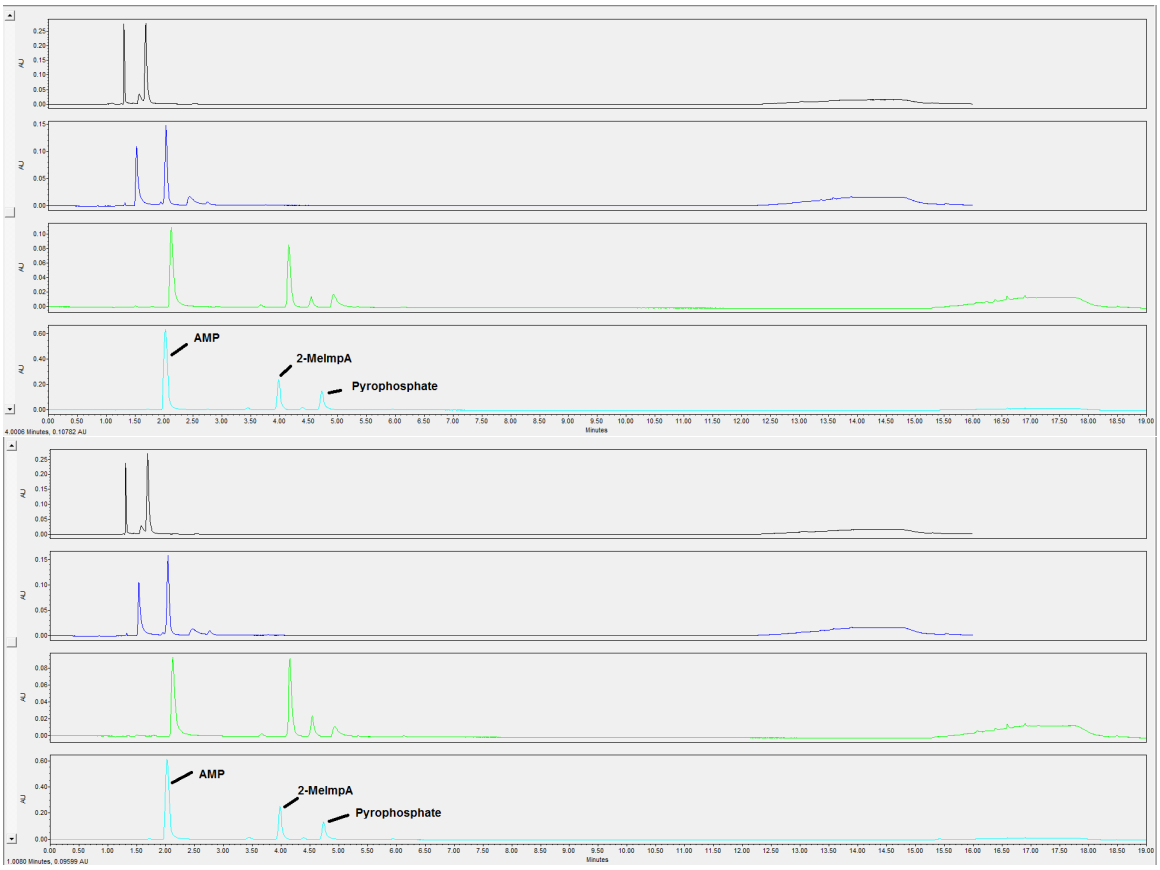


Figure 28: Sample C2 over time. Top four chromatograms 1 mM ATP + 5 mM 2-MeImpA. Bottom chromatogram 10 mM PDADMAC + 1 mM ATP + 5 mM 2-MeImpA. First measurement after 4 days, second after 10 days, third after 12 days, and fourth after 21 days.

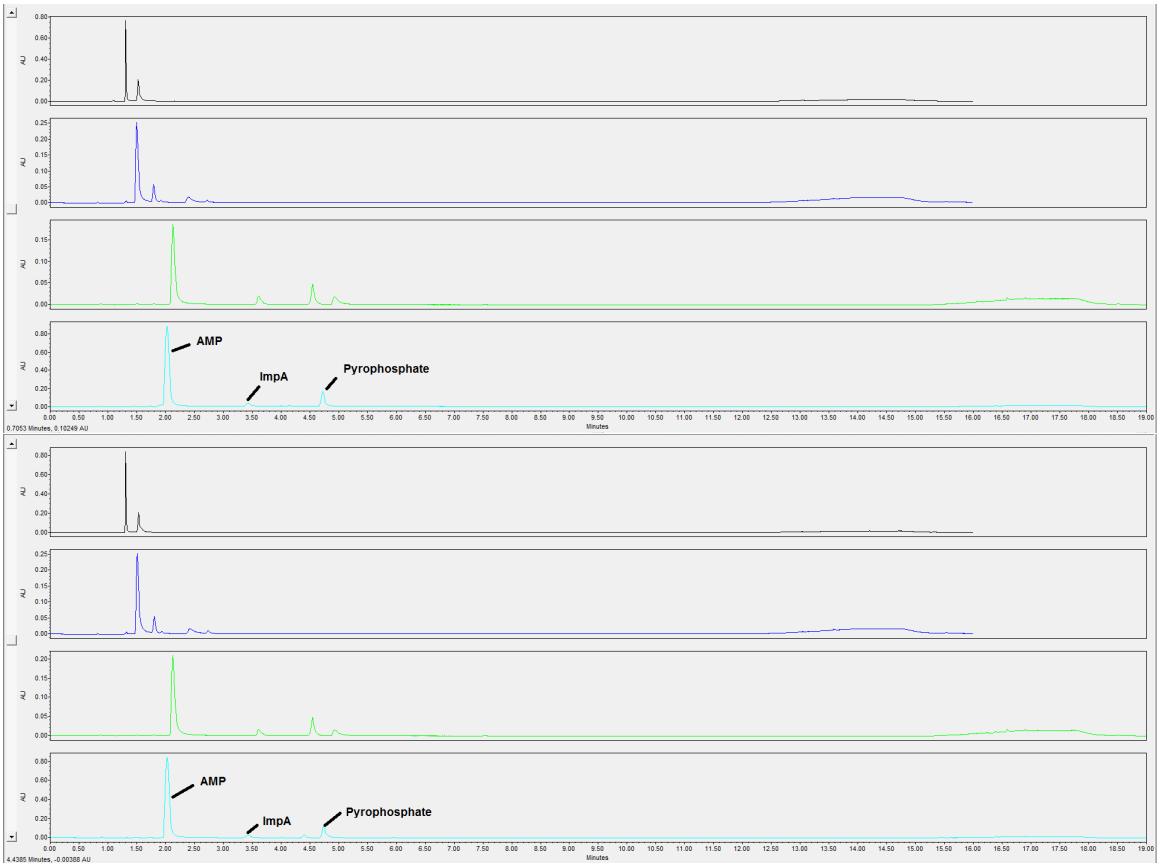


Figure 29: Sample D1 over time. Top four chromatograms 2.5 mM ATP + 5 mM ImpA. Bottom chromatogram 10 mM PDADMAC + 2.5 mM ATP + 5 mM ImpA. First measurement after 4 days, second after 10 days, third after 12 days, and fourth after 21 days.

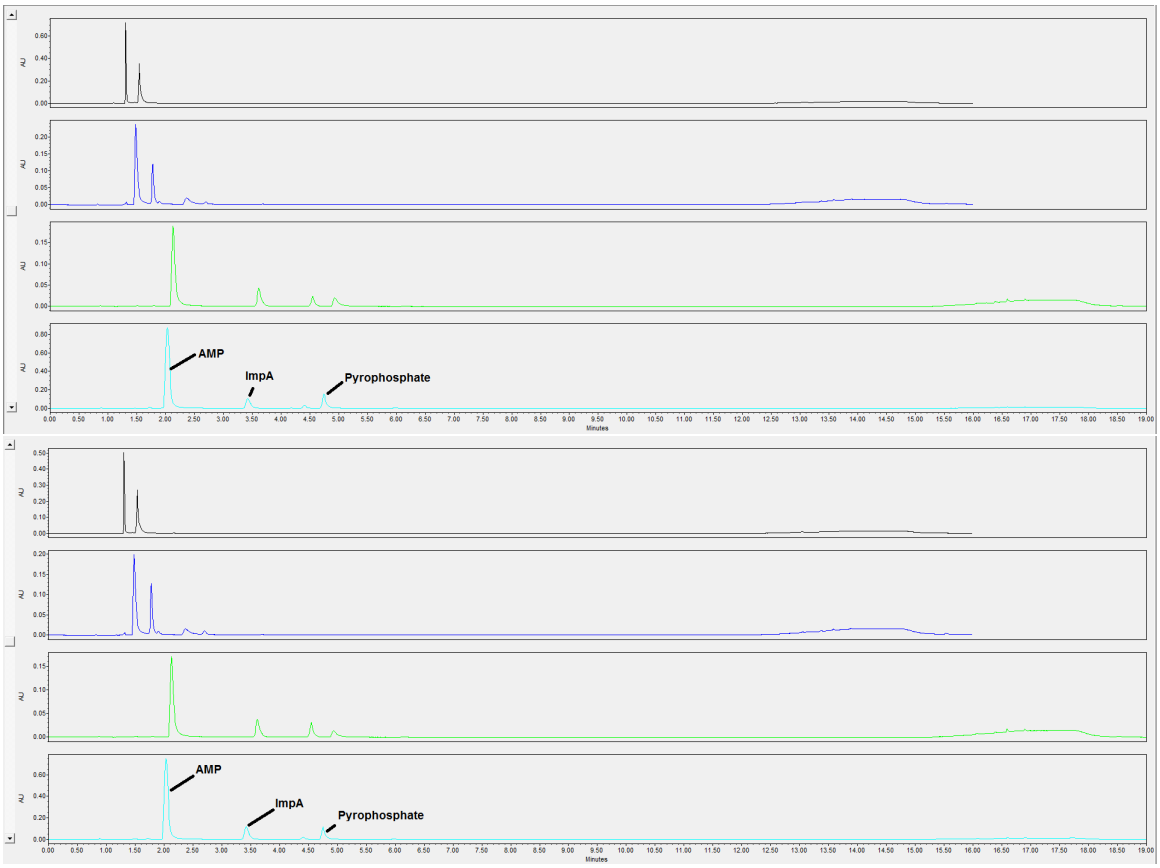


Figure 30: Sample D2 over time. Top four chromatograms 1 mM ATP + 5 mM ImpA. Bottom chromatogram 10 mM PDADMAC + 1 mM ATP + 5 mM ImpA. First measurement after 4 days, second after 10 days, third after 12 days, and fourth after 21 days.

9 Appendix C

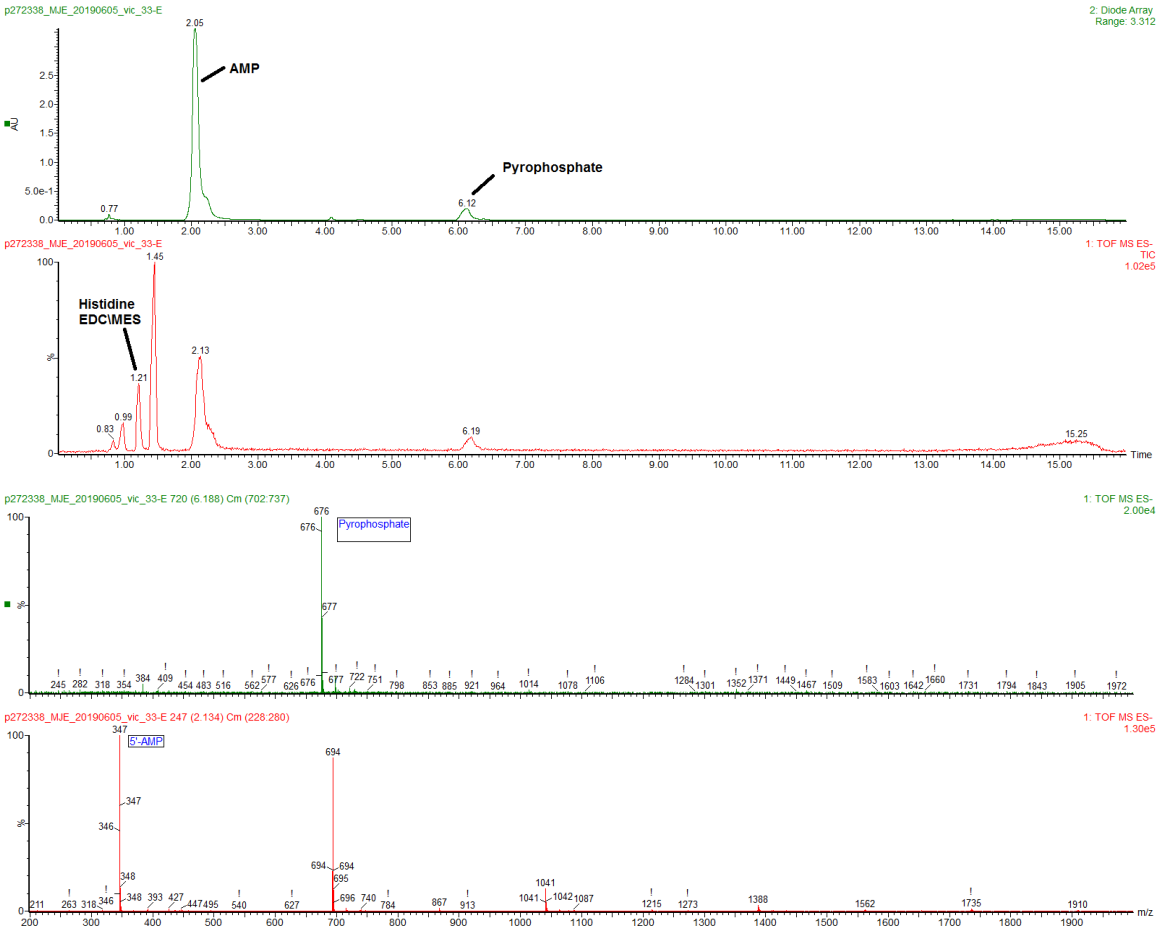


Figure 31: Mass spectrometry analysis to identify peaks at different retention times for the UPLC analysis of the samples below. Sample initially contained 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5

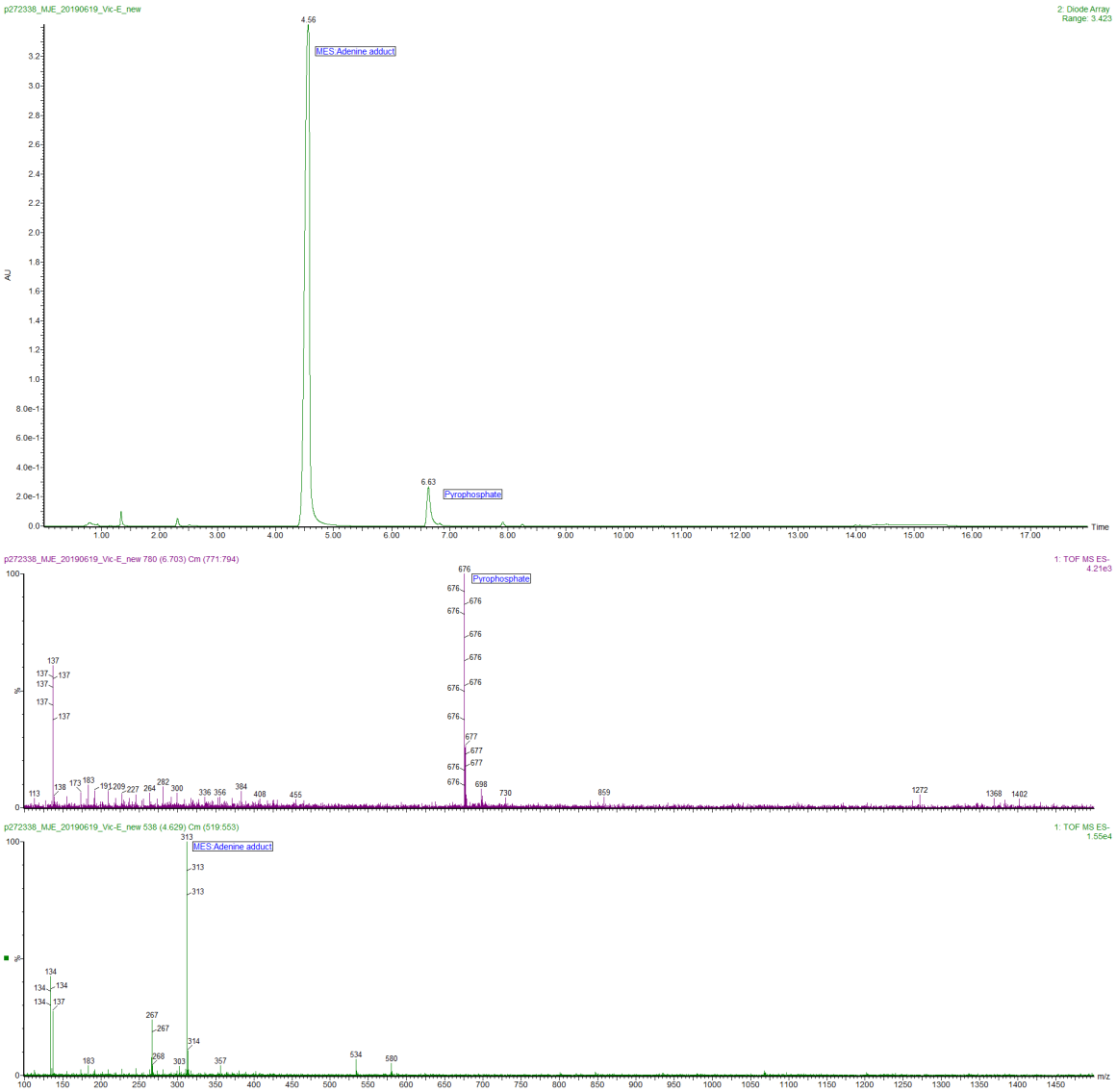


Figure 32: Mass spectrometry analysis of anomalous sample E. The peak at t= 4min corresponds to the proposed MES:adenine adduct and the peak at t= 6min corresponds to pyrophosphate. Sample initially contained 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

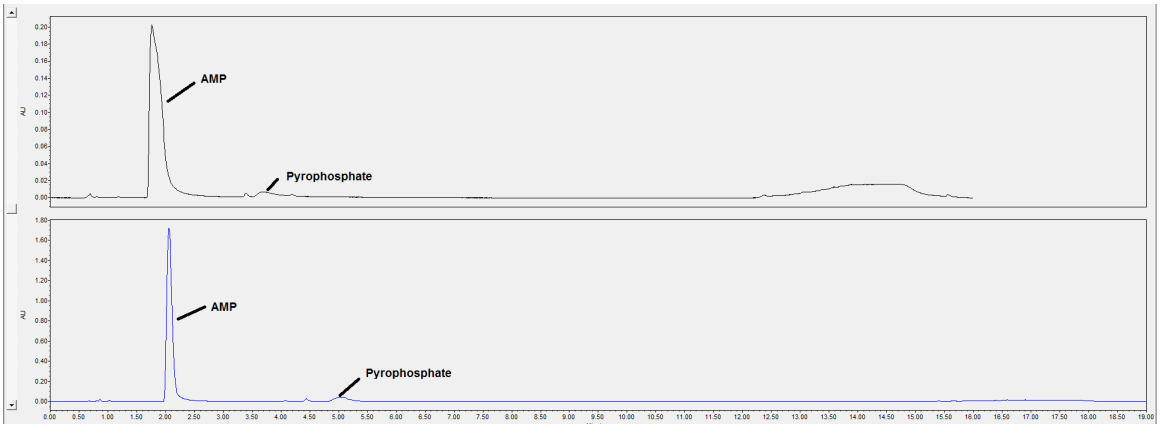


Figure 33: Sample A; first measurement after 5 days; second measurement after 16 days. Sample initially contained 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

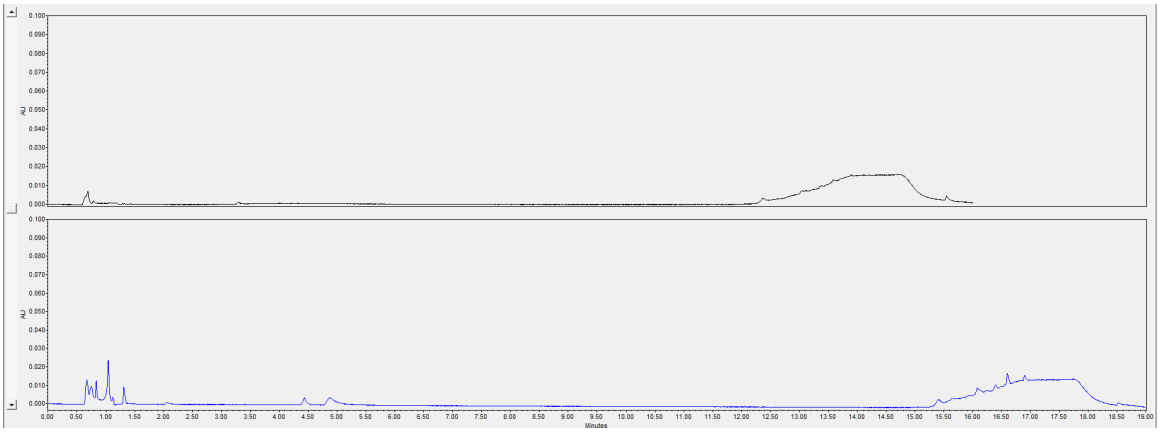


Figure 34: Sample B; first measurement after 5 days; second measurement after 16 days. No peak identified due to absence of absorbing compounds at 254nm. Sample initially contained poly-L-histidine (25 mM), and EDC (100mM) at pH 5.

Figure 35: Sample C; first measurement after 5 days; second measurement after 16 days. Sample initially contained 5'-AMP (11 mM), and EDC (100mM) at pH 5.

Figure 36: Sample D; first measurement after 5 days; second measurement after 16 days. Sample initially contained 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

Figure 37: Sample E; first measurement after 5 days; second measurement after 16 days. Interesting behaviour since the peak at t=2min has disappeared slowly and was replaced by a peak at t=4.2min. Sample initially contained 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.

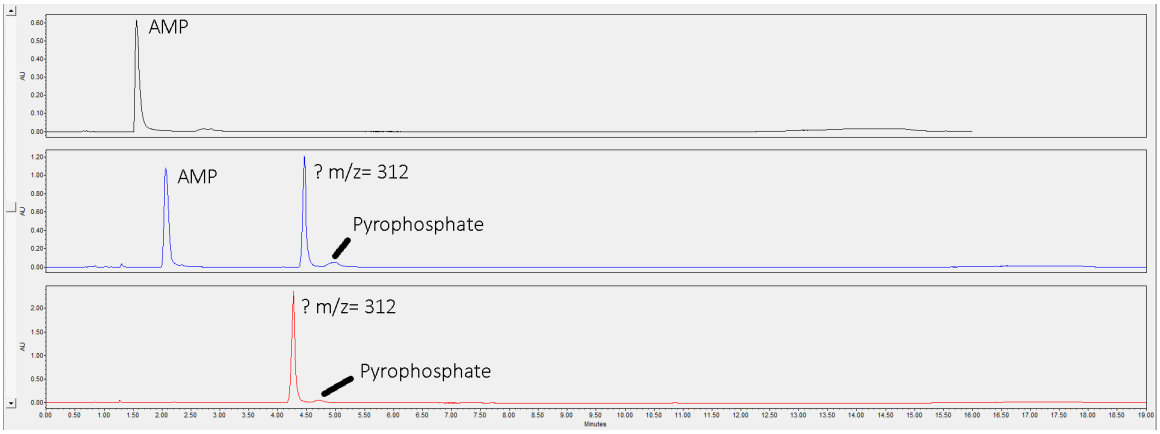


Figure 38: Sample E; first measurement after 5 days; second measurement after 16 days, and third measurement after 21 days. Interesting behaviour since the peak at t=2min has disappeared slowly and was replaced by a peak at t=4.2min. Sample initially contained 5'-AMP (11 mM), poly-L-histidine (25 mM) and EDC (100mM) at pH 5.