# $^1$H and $^{31}$P NMR study of ADP and GDP spontaneous synthesis from 2-methylimidazole-activated AMP and GMP nucleotides and inorganic phosphate

**José Pereira * and Ana Cadete**

*CEQUP/Instituto de Ciências Biomédicas de Abel Salazar, University of Porto, Largo Professor Abel Salazar, 2, 4099-003 PORTO, Portugal*

***Received (in Cambridge) 2nd February 1999, Accepted 19th March 1999***

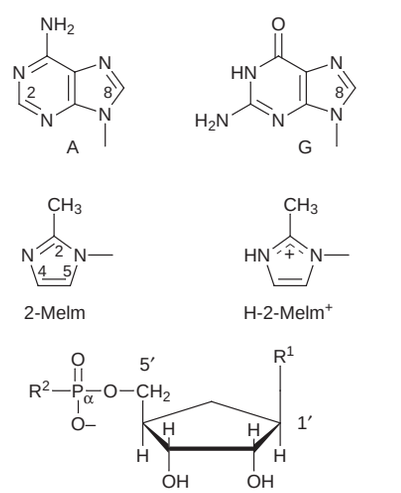
$^1$H and $^{31}$P NMR data show that adenosine 5′-(2-methylimidazol-1-ylphosphonate) (2-MeImpA) and guanosine 5′-(2-methylimidazol-1-ylphosphonate) (2-MeImpG), known as possible prebiotic precursors of polynucleotides, produce the corresponding diphosphonucleotides in sodium phosphate solution at pD 7.6. Phosphate ions also enhance the hydrolysis of these molecules by activating a water molecular associated as a weak complex with the phosphoryl moiety of 2-MeImpA and 2-MeImpG. A kinetic study was done by quantitative $^1$H NMR spectroscopy and mechanistic hypotheses were tested by semiempirical PM3 modelling.

## Introduction

Imidazole-activated nucleotides can be formed in simulated prebiotic conditions and have been used as precursors for *in vitro* synthesis of various kinds of oligoribonucleotides[1–4] and oligodeoxyribonucleotides[5] using as templates pre-existent polynucleotides. Other lines of work study the effectiveness of inorganic catalysts or templates, more likely to be present in prebiotic conditions, like metal ions[6] or montmorillonite clays and other minerals.[7] Detailed mechanistic and kinetic analysis of nucleotide imidazolide hydrolysis and polymerisation is a fundamental step to understand the action of these particularly reactive oligonucleotide precursors. The general reaction of polymerisation occurs through a nucleophilic substitution on the phosphoryl moiety of the activated nucleotide by another nucleotide hydroxy group, imidazole being the leaving group. The $Mg^{II}$ ion has an important role in the process,[8] further activating the phosphoryl moiety. Depending on the conditions (templates, catalysts, pH, temperature) several kinds and lengths of oligonucleotides may result, even dimers like NppN (pyrophosphoryl bridge between two nucleotides). The activated phosphoryl group is in fact sensitive to any nucleophile, suggesting a method for the prebiotic synthesis of nucleotide polyphosphates by reaction with inorganic phosphate or pyrophosphate in solution. In this work, the effect of phosphate concentration on 2-methylimidazole-activated nucleotides was studied in aqueous systems containing sodium phosphate and adenosine 5′-(2-methylimidazol-1-ylphosphonate) (2-MeImpA) or guanosine 5′(2-methylimidazol-1-ylphosphonate) (2-MeImpG) (Scheme 1).†

NMR was chosen as the qualitative and quantitative analytical technique in this study as an alternative to the commonly used HPLC because it provides an equally accurate and more straightforward experimental method. It also has the advantage of providing structural insight into the reaction products. The pD value of 7.6 was chosen for the experiments because it is low enough to avoid reaction of 2-MeImpN with $OH^-$ and large enough to limit the reaction rate so that NMR spectra acquisition times are negligible. It is also inside the presumed pH values of prebiotic sea water.

† In the text 2-MeImpA stands for 2-MeImpA$^-$ and H-2-MeImpA$^\pm$, 2-MeImpG stands for 2-MeImpG$^-$ and H-2-MeImpG$^\pm$, 2-MeImpN stands for 2-MeImpA and 2-MeImpG. The word phosphate or the symbol $P_i$ designate $H_2PO_4^-$ and $HPO_4^{2-}$.




| | |
|---|---|
| 2-MeImpA$^-$ | $R^1$ = A, $R^2$ = 2-MeIm |
| H-2-MeImpA$^\pm$ | $R^1$ = A, $R^2$ = 2-MeIm$^+$ |
| 2-MeImpG$^-$ | $R^1$ = G, $R^2$ = 2-MeIm |
| H-2-MeImpG$^\pm$ | $R^1$ = G, $R^2$ = 2-MeIm$^+$ |
| 2-MeImpR$^-$ | $R^1$ = –$NH_2$, $R^2$ = 2-MeIm |
| H-2-MeImpR$^\pm$ | $R^1$ = –$NH_2$, $R^2$ = 2-MeIm$^+$ |

**Scheme 1**

## Results and discussion

### Analysis and interpretation of 2-MeImpG and 2-MeImpA hydrolysis NMR spectra

Typical $^1$H and $^{31}$P spectra of 2-MeImpG and 2-MeImpA in phosphate buffer are presented in Fig. 1 after some of the reaction has occurred. The spectra taken in 0.2 M TRIS buffer instead of phosphate buffer are identical to the spectra in Fig. 1, but the signals assigned to ADP or GDP species are absent.

The chemical shifts presented in Fig. 1 are constant for all experiments reported in this work with respect to buffer type, buffer concentration and time and so correspond to a full $^1$H and $^{31}$P NMR spectral characterisation of 2-MeImpA and 2-MeImpG and their hydrolysis products, under the standard conditions used in all experiments: 0.02 M 2-2-MeImpN, 0.5 M NaCl, pD 7.6, 308 K (see Experimental).

The assignment of the 2-MeImpG spectrum was performed by comparing with spectra of GMP, guanosine and

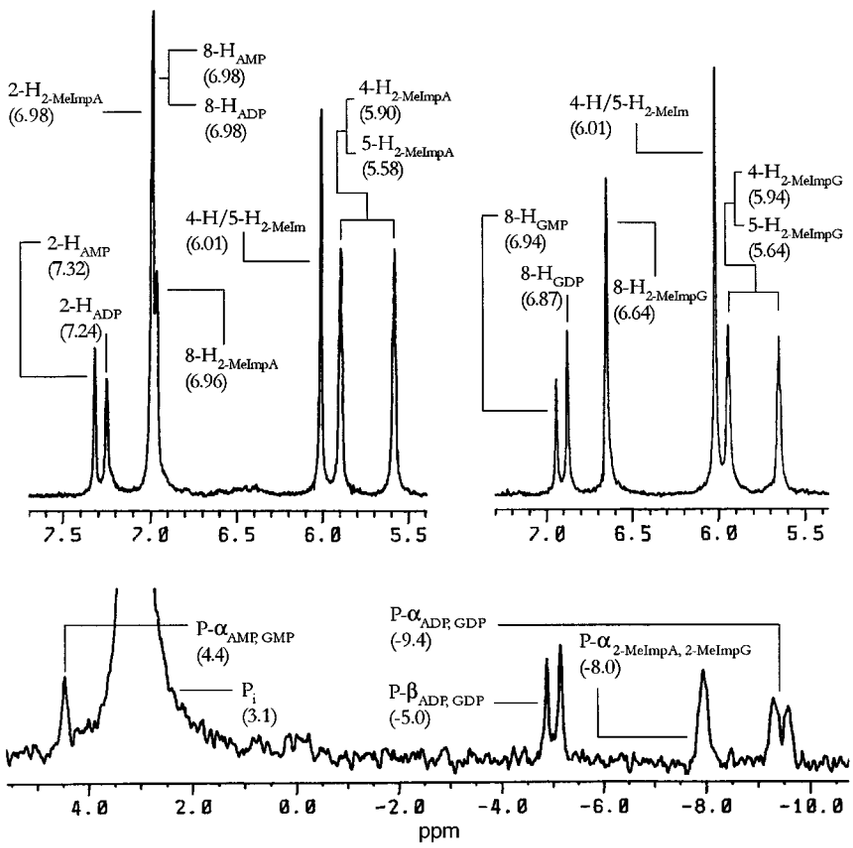



**Fig. 1** Chemical shifts are in ppm. (Top) Typical 2-MeImpA (left) and 2-MeImpG (right) $^1$H NMR spectra (aromatic region) in phosphate buffer, showing reactants and products signals. Shown spectra are after approximately 4 h of reaction, $[P_i] = 0.20$ M in 2-MeImpA solution and $[P_i] = 0.40$ M in 2-MeImpG solution; all other conditions are standard (see Experimental). The methyl signals $Me_{2\text{-MeIm}}$ (1.36 ppm), $Me_{2\text{-MeImpA}}$ (1.14 ppm) and $Me_{2\text{-MeImpG}}$ (1.20 ppm) are not shown. (Bottom) Typical $^{31}$P NMR spectra for 2-MeImpA or 2-MeImpG samples in phosphate buffer. Shown spectrum is for 2-MeImpG solution after approximately 13 h of reaction, $[P_i] = 0.20$ M; all other conditions are standard (see Experimental). When 0.2 M TRIS buffer is used instead of phosphate buffer all the signals labelled ADP or GDP are not present; the other signal chemical shifts are the same as for phosphate buffer solutions.

2-methylimidazole taken under the standard conditions used in this work. The suspicion that GDP was present in 2-MeImpG phosphate buffer solutions stems mainly from analysis of the $^{31}$P NMR spectrum, which shows two doublets that are not present when TRIS buffer is used. These correspond to two phosphorus nuclei that are mutually coupled, as can be proved by a simple $^{31}$P–$^{31}$P COSY experiment, with an apparent coupling constant of 32 Hz. One of these doublets also shows an unresolved fine structure which was attributed by a $^1$H–$^{31}$P heteronuclear correlation experiment to coupling with the H-5′ nuclei from the ribosyl moiety. This observation alone excludes the possibility of GppG being one of the hydrolysis products and strongly suggests the presence of GDP in solution, as well as the expected GMP. In fact, the GDP $^{31}$P NMR spectrum was found to match perfectly the two doublets observed in the 2-MeImpG $^{31}$P NMR spectrum in phosphate buffer. The $^1$H NMR spectrum of 2-MeImpG in phosphate buffer also shows an unexpected signal for 8-H which was found to have the same chemical shift as the 8-H signal in the GDP $^1$H NMR spectrum, thus corroborating the conclusions drawn from the $^{31}$P NMR spectral analysis.

To qualitatively analyse the 2-MeImpA reactions all the experiments mentioned above for 2-MeImpG were reproduced. To reach the assignments presented in Fig. 1 (left and bottom) the AMP and ADP spectra were compared with the 2-MeImpA spectra. The 2-MeImpA $^{31}$P NMR spectra in phosphate buffer show exactly the same pattern of signals present in the 2-MeImpG $^{31}$P NMR spectra and so the reaction products are AMP and ADP if the 2-MeImpA hydrolysis is performed in phosphate buffer and only AMP if TRIS buffer is used instead.

The spontaneous phosphorylation of 2-methylimidazolide-activated nucleotides from inorganic phosphate has not previously been reported and the kinetic and mechanistic aspects of this reaction are the subjects of the next sections. No evidence for nucleotide dimers or polymers was found in the spectra.

## Kinetic model

The NMR integrals are proportional to concentrations and not to activities and so all derivations will be made using concentrations. In view of the NMR species analysis, eqns. (1) and (2)

$$\text{2-MeImpN} \xrightarrow{k_1'} \text{2-MeIm} + \text{NMP} \qquad (1)$$

$$\text{2-MeImpN} + P_i \xrightarrow{k_2'} \text{2-MeIm} + \text{NDP} \qquad (2)$$

are necessary to describe the 2-MeImpN system for the pH value of 7.6 used in this work.

Phosphate concentration will always be considered in great excess over 2-MeImpN concentration and therefore included in the rate coefficients when necessary. By making the simplest mechanistic assumptions the main rate equations applicable to the degradation of 2-MeImpN are given in eqns. (3)–(5).

$$-\mathrm{d[2MeImpN]}/\mathrm{d}t = k_{obs}[\text{2-MeImpN}] \qquad (3)$$

$$\mathrm{d[NMP]}/\mathrm{d}t = k_1'[\text{2-MeImpN}] \qquad (4)$$

$$\mathrm{d[NDP]}/\mathrm{d}t = k_2'[\text{2-MeImpN}] \qquad (5)$$

Based on eqns. (1)–(5), expressions for the time dependence of [2-MeImpN], [2-MeIm], [NMP] and [NDP] can be derived [eqns. (6)–(9)].

**Table 1** Observed rate coefficients and final ratio between NMP and NDP mole fractions in solution of 0.02 M 2-MeImpA and 2-MeImpG; pD 7.6, 308 K

| | 2-MeImpA | | | | | 2-MeImpG | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $[P_i]$/M | 0[a] | 0.20 | 0.30 | 0.40 | 0.50 | 0[a] | 0.20 | 0.30 | 0.40 | 0.50 |
| $k_{obs}/h^{-1}$ | 0.036 | 0.083 | 0.116 | 0.152 | 0.180 | 0.027 | 0.071 | 0.103 | 0.132 | 0.165 |
| $\chi_{NMP}:\chi_{NDP}$ | — | 0.78 | 0.71 | 0.69 | 0.63 | — | 0.54 | 0.50 | 0.45 | 0.43 |

[a] 0.20 M TRIS buffer.

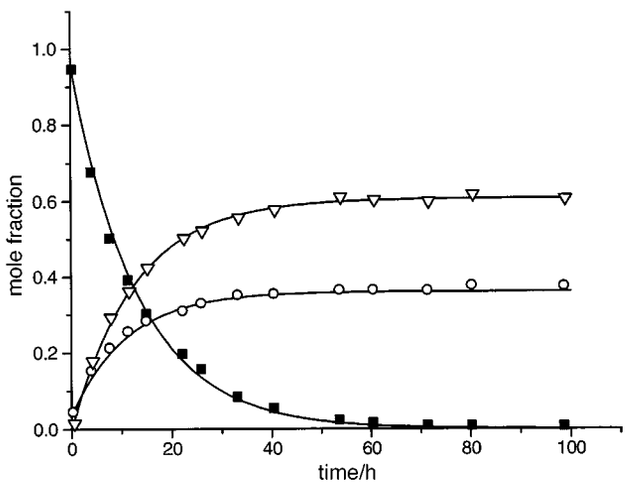



**Fig. 2** Time evolution of [2-MeImpG] (squares), [GMP] (circles) and [GDP] (triangles) in 0.20 M phosphate buffer as monitored by the integrals of 8-$H_{2\text{-MeImpG}}$, 8-$H_{GMP}$ and 8-$H_{GDP}$ NMR signals. Actual initial concentration of 2-MeImpG was 0.019 M due to time zero hydrolysis. Fittings from eqns. (6), (8) and (9).

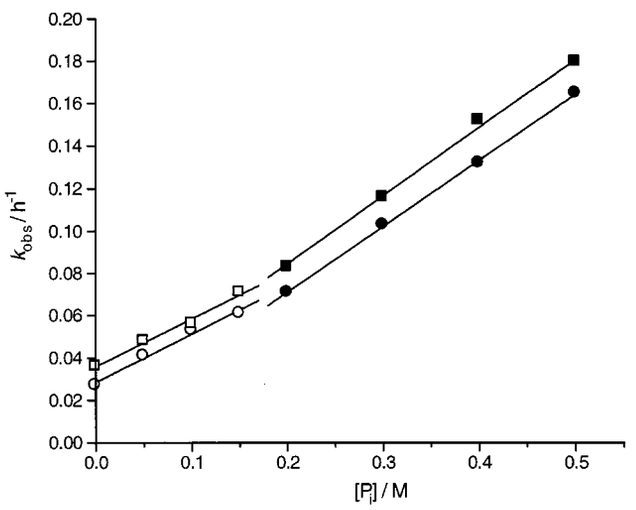



**Fig. 3** Effect of phosphate concentration on observed rate coefficients. Squares are for 2-MeImpA and circles are for 2-MeImpG (0.002 M solutions). Open symbols represent values obtained from mixed buffer solutions with [TRIS]:[phosphate] equal to 0.20:0, 0.15:0.05, 0.10:0.10 and 0.05:0.15 M.

$$[\text{2-MeImpN}] = [\text{2-MeImpN}]_0\, e^{-tk_{obs}} \quad (6)$$

$$[\text{2-MeIm}] = [\text{2-MeImpN}]_0(1 - e^{-tk_{obs}}) \quad (7)$$

$$[\text{NMP}] = (k_1'/k_{obs})[\text{2-MeImpN}]_0(1 - e^{-tk_{obs}}) \quad (8)$$

$$[\text{NDP}] = (k_2'/k_{obs})[\text{2-MeImpN}]_0(1 - e^{-tk_{obs}}) \quad (9)$$

The coefficients $k_1'$ and $k_2'$ can be calculated indirectly from $k_{obs}$, [NMP] and [NDP] using eqns. (10) and (11).

$$k_1' = k_{obs}/(1 + ([\text{NDP}]/[\text{NMP}])) \quad (10)$$

$$k_2' = k_{obs}/(1 + ([\text{NMP}]/[\text{NDP}])) \quad (11)$$

### NMR data, $k_{obs}$ calculation and the buffer effect

Fig. 2 shows a typical graph of species mole fraction evolution with time. Some hydrolysis is noticeable at time zero and was accounted for in the fittings by letting $[\text{2-MeImpN}]_0$ be an adjustable parameter in eqns. (6), (7) and (8). The long term storage and occasional opening of the vacuum containers where the 2-MeImpA and 2-MeImpG solids were stored may have caused some water vapor to be adsorbed leading to some hydrolysis.

Very good agreement was generally found between data and the exponential functions for all the $^1H$ nuclei, but the $k_{obs}$ values obtained from the signals 4-H/5-$H_{2\text{-MeIm}}$ and 4-$H_{2\text{-MeImpG}}$, which overlap at the base, had to be excluded from the average $k_{obs}$ calculations. For 2-MeImpA solutions, only the overlapping signal sets of 2-H and 8-H were not used. The $^{31}P$ signals could not be used for quantitative measurements due to unacceptable spectral noise.

The calculated $k_{obs}$ values for phosphate buffer only solutions are presented in Table 1 and were obtained by fitting eqns. (6) to (9) to the NMR mole fraction data and averaging for the selected signals. The marked effect of phosphate concentration on $k_{obs}$ is obvious for both nucleotides. The standard deviation associated with these values is smaller than 5% of the average and reveals the unavoidable errors on integral measurements mostly due to phase and baseline correction of spectra.

Solutions with mixed TRIS and sodium phosphate buffer were analysed to investigate the effect of buffers on the experimental rate constants. In Fig. 3 the first four data points were taken from solutions with mixed buffer and show a different apparent linear slope than the last four data points that correspond to phosphate buffer only solutions. TRIS seems to have an increasing effect on $k_{obs}$ relative to the expected linear extrapolation of the four last points to zero concentration and so the inertness of this buffer when used for kinetic calculations of this kind of compound on aqueous solution can be questioned. A possible mechanism for the influence of TRIS and phosphate buffers on 2-MeImpN hydrolysis rate will be discussed in the next sections.

The overall observation is that the reaction rates increase with phosphate concentration and that 2-MeImpA reactions are slightly faster. The different contributions to the rates can be analysed using the final mole fractions of NMP and NDP in solutions with varying concentrations of phosphate buffer.

### Rate coefficients for NMP and NDP formation

The $\chi_{NMP}:\chi_{NDP}$ values in Table 1 are the average of the 2-$H_{AMP}$ and 2-$H_{ADP}$ integral ratio measured for 2-MeImpA solutions and for 8-$H_{GMP}$ and 8-$H_{GDP}$ integral ratio measured for 2-MeImpG solutions at the end of the reaction. The phosphate concentration dependence of $k_1'$ and $k_2'$ values, derived from data in Table 1 and eqns. (10) and (11), is presented in Fig. 4 for $k_1'$ and Fig. 5 for $k_2'$ (logarithmic form).

The $k_1'$ dependence on phosphate concentration was unexpected in view of the proposed model [eqn. (1)] but it is in fact a consequence of the observed relative increase of $k_{obs}$ over $\chi_{NDP}:\chi_{NMP}$, that is, $\chi_{NMP}$ also responds positively to the

**Table 2** Calculated rate coefficients for eqns. (1) and (2)[a]

| | $k_1$[b]/h$^{-1}$ | $k_{1p}$[b]/M$^{-1}$ h$^{-1}$ | $k_2$[c]/M$^{-1}$ h$^{-1}$ | $n$[c] |
|---|---|---|---|---|
| 2-MeImpA | 0.014 ± 0.003 | 0.11 ± 0.01 | 0.21 ± 0.01 | 0.94 ± 0.02 |
| 2-MeImpG | 0.009 ± 0.001 | 0.081 ± 0.003 | 0.23 ± 0.01 | 0.99 ± 0.02 |

[a] The error intervals presented are the linear fitting parameters standard errors. [b] From fitting of eqn. (12), $n$ set to 1. [c] From the fitting of eqn. (13) logarithmic form.

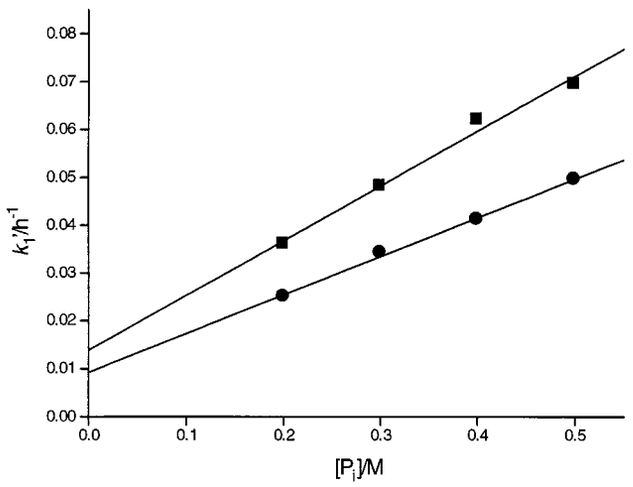



**Fig. 4** Calculation of $k_1$ and $k_{1p}$ by fitting the eqn. (12) to $k_1'$ dependence on phosphate concentration, $n$ set to 1. Squares are for 2-MeImpA and circles are for MeImpG (0.02 M solutions).

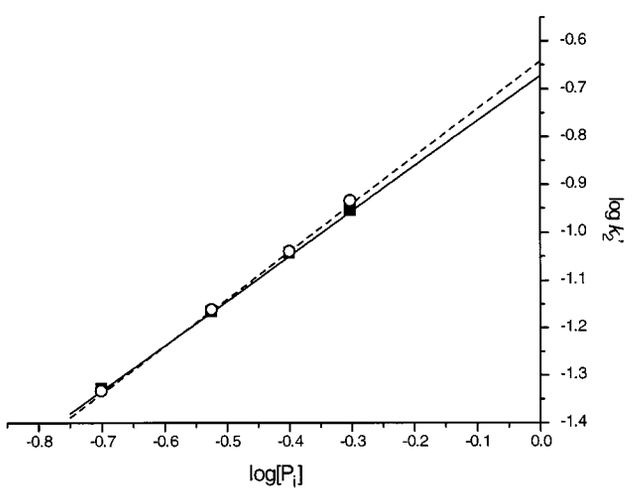



**Fig. 5** Calculation of $k_2$ by fitting the logarithmic form of eqn. (13) to $k_2'$ dependence on phosphate concentration. Squares and solid line are for 2-MeImpA, circles and dashed line are for 2-MeImpG (0.02 M solutions).

increasing concentration of phosphate. This observation was interpreted as a result of phosphate catalysed 2-MeImp hydrolysis and not caused by NDP species hydrolysis, which has reasonable kinetic stability at pH 7.6. This required therefore the definition of $k_1'$ as in eqn. (12), where $k_1$ is the hydrolysis

$$k_1' = k_1 + k_{1p}[P_i]^n \quad (12)$$

rate coefficient and $k_{1p}$ is the phosphate catalysed hydrolysis rate coefficient. For $k_2'$ the phosphate dependence can be described as in eqn. (13).

$$k_2' = k_2[P_i]^n \quad (13)$$

Fig. 4 and 5 show in graphical form the $k_1'$ and $k_2'$ data used to calculate $k_1$, $k_{1p}$ and $k_2$ by fitting eqn. (12), with $n$ equal to 1, and the logarithmic form of eqn. (13); Table 2 summarises the calculated values. Eqn. (12) was used with $n$ equal to 1 because the data are apparently linear. This linearity, observed for both nucleotide derivatives, indicates a first order dependence on phosphate concentration for the phosphate catalysed hydrolysis reaction [embedded in eqn. (1)]. This suggests a process in which a phosphate ion activates a water molecule by hydrogen bonding leaving more net negative charge in the water oxygen atom that attacks more effectively the positively charged phosphorus atom of the nucleotide. This might also be the mechanism by which TRIS enhances the 2-MeImpN hydrolysis, acting as a base for one of the hydrogen atoms of a solvation water molecule.

The fitting of the logarithmic form of eqn. (13) to the $k_2'$ data yielded a value of approximately 1 for $n$ on both nucleotide derivatives (Table 2), which is consistent with a nucleophile substitution of second order ($S_N2$) for the mechanism associated with eqn. (2).

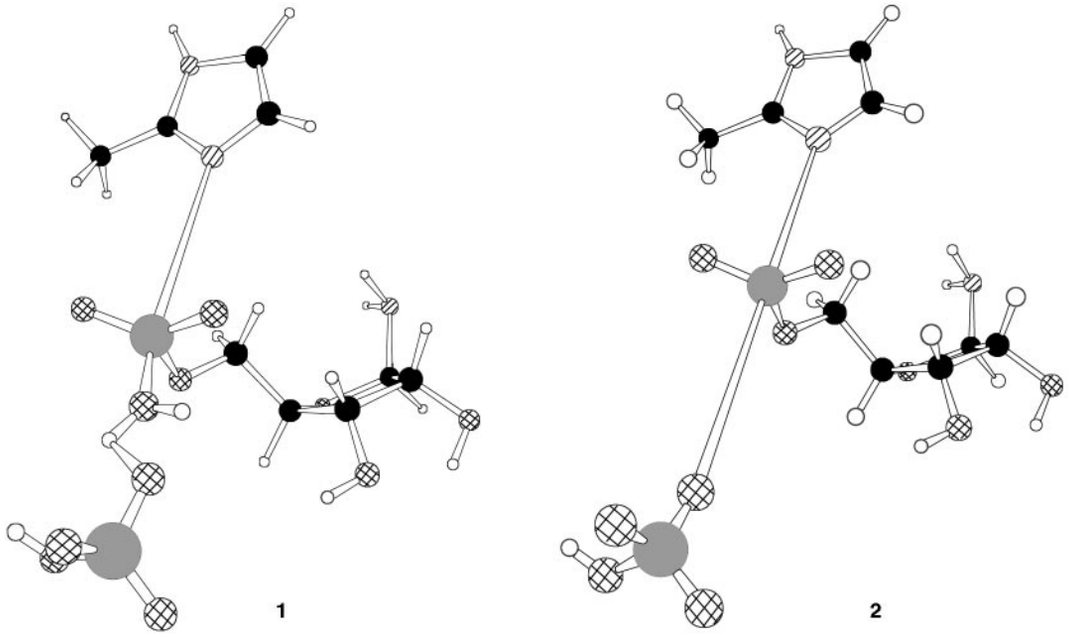
The difference in $k_{obs}$ observed for 2-MeImpA and 2-MeImpG is in fact a consequence of a slightly faster rate of hydrolysis (catalysed and non-catalysed) for 2-MeImpA, the rate coefficient for the formation of ADP and GDP being very similar in both nucleotides. There weren't expected to be large differences between the reactivity of the two activated nucleotides given the bond distance between the reaction location (the nucleotide phosphoryl) and the nucleotide base. As will be discussed below, the dependence of $k_1$ on pD can account for the observed rate coefficient differences between 2-MeImpA and 2-MeImpG because its p$K_a$ values can be slightly different.

At pD 7.6 both the protonated and deprotonated forms of 2-MeImpN exist in solution in approximately equal amount (p$K_a$ ~ 7.74)[8] but the protonated form, H-2-MeImpG$^{\pm}$, reacts much more rapidly with water, due to the positive charge on the 2-methylimidazole ring. The calculated rate coefficient $k_1$ should then be about half of the maximum possible value. Allowing for this, for the fact that these experiments were done in $D_2O$ and not $H_2O$, and also for experimental error, a value of two times 0.009 h$^{-1}$ for 2-MeImpG $k_1$ is still less than the previously published value[8] of 0.033 h$^{-1}$. The $k_1$ values calculated in this work are in fact for zero concentration of buffer, which could account for the difference observed. The rate coefficients $k_{1p}$ and $k_2$ have not previously been reported and also depend on pD not only because of the 2-methylimidazole moiety protonation but also because of phosphate protonation. Two species of phosphate are in solution at pD 7.6 with different charges ($HPO_4^{2-}$ and $H_2PO_4^-$) and probably with different reactivities. Nevertheless, given the $k_2$ values here determined and the rate coefficients determined for template directed dimerisation of 2-methylimidazole activated nucleotides,[9] inorganic phosphate, if present in solution in equivalent amount to the nucleotide and at pH values near 7.6, is able to compete effectively with template directed polymerisation.

### Molecular modelling and reaction simulation

Two different reactions are proposed for 2-MeImpN in phosphate buffer: the hydrolysis [eqn. (1)] and the 2-methylimidazole substitution by phosphate [eqn. (2)]. We propose also that hydrolysis can be accelerated by phosphate, or other bases, like TRIS, through a mechanism in which a solvation water molecule associated with P-α, structure **1** in Scheme 2, becomes more nucleophilic when phosphate, for example, forms a hydrogen bond with one of the hydrogen atoms of the associated water molecule.

Computational modelling of the reactions was performed to validate the proposed mechanisms and to assess the importance of protonation on the imidazole and on the phosphate ion. Based on the experimental results described so far a semi-empirical PM3 modelling of reagent association as "weak complexes", of which two examples are shown in Scheme 2,




**Scheme 2**

was performed. The representation of exchangeable hydrogen atoms as $^1H$ or $^2H$ didn't affect the results of the calculations reported below. To speed up the calculations, and since the results for the two nucleotide derivatives were qualitatively equal and quantitatively very similar, a smaller model was used in which the purine base moiety was replaced by an $NH_2$ group, molecules 2-MeImpR$^-$ or H-2-MeImpR$^\pm$ (Scheme 1).

In the PM3 equilibrium geometries of H-2-MeImpR$^\pm$ and 2-MeImpR$^-$ the phosphorus atom has a clearly positive charge but more so on H-2-MeImpR$^\pm$ (electrostatic fit charge of about +2.5). In this molecule the electronic density in the P–N bond (of approximately 2 Å in length) is less than in 2-MeImpR$^-$, consequently being easier to break.

PM3 energy minimisations show that water can form a weak complex with 2-MeImpR$^-$ and H-2-MeImpR$^\pm$, stabilised by a dative bond to the P–N σ* MO and one or two hydrogen bonds as in Scheme 2, complex **1**. Hydrolysis, and the consequent P–N bond breakage, can proceed *via* nucleophilic attack at one of the hydrogen atoms on the associated water molecule. This kind of relayed attack on the phosphoryl group was simulated by associating this complex with a $HPO_4^{2-}$ or $H_2PO_4^-$ ion as shown in **1**. These "complexes" were then subjected to geometry optimisation (energy minimisation) without constraints and with several initial distances along the O–H–O direction between the reactants. Above some critical distance the energy minimisation process physically separates the reactants but below that distance the energy minimisation procedure reproduces the actual reaction, breaking the P–N bond and generating the ribose 1-amino 5-phosphate.

The attack of $H_2PO_4^-$ and $HPO_4^{2-}$ on H-2-MeImpR$^\pm$ and 2-MeImpR$^-$ was simulated in a similar way. Weak complexes were built in which the phosphate attacks the phosphoryl group along a straight line that contains the atoms (N–P)–(O–P), as in **2**, Scheme 2. Products were 2-methylimidazole and ribose 1-amino 5-diphosphate.

These critical distances, obtained in the simulated gas phase, lie between 1.8 to 5.5 Å and relate to transition states inasmuch as the selected intrinsic reaction co-ordinates are concerned. Two such possible transition structures determined in this work in the gas phase are shown in Scheme 2. Water solvation tends to stabilise localised charge (ions) and to shield the reactants interaction, thus increasing the energy barrier and slowing down the reactions. However, the proposed reaction mechanisms are plausible also in aqueous solution, possibly with lower critical distances. Taking these distances as a crude indication of reactivity, H-2-MeImpN$^\pm$ and $HPO_4^{2-}$ are the most effective reactants (longer critical distance) and should account for the most part of the experimental rate coefficients.

## Conclusions

For relatively simple systems like the ones studied in this work NMR is a good qualitative and quantitative tool. Overlapping of signals can, however, hinder the study of more complex systems, namely, systems where oligonucleotides are a product. In that situation it would be difficult to distinguish between different kinds and sizes of polymers. The choice of buffers is somewhat limited for NMR experiments since it is necessary to avoid overlapping of buffer signals with the subject substance signals, but direct influence of the buffer on the measured coefficient rates is also an issue independent of the chosen analytical technique. The two buffers used in this work were shown to have a non-negligible effect on the 2-MeImpA and 2-MeImpG hydrolysis rates. Phosphate catalyses 2-MeImpN hydrolysis by activating a solvation water molecule and TRIS also clearly increases the 2-MeImpN hydrolysis rate coefficients (Fig. 3) probably by the same mechanism, acting as a base.

Optimistic and pessimistic arguments regarding chemical replication in prebiotic conditions have at this moment a relative weight that is difficult to evaluate [10] but for any hypothesis to be valuable robustness is an important attribute. As was mentioned in the Introduction the ribonucleoside 5′-(2-methylimidazol-1-ylphosphonates) can polymerise in a template, be it a previously available polyribonucleotide or an inorganic surface like a montmorillonite clay, without the need for a catalyst other than the template itself. But, as was shown in this work, the hydrolyses of these activated nucleotides can be easily enhanced by the presence in solution of phosphate ions, and probably by other bases. The observed 2-methylimidazole nucleophilic substitution by phosphate is also very effective at pD 7.6. 2-MeImpA and 2-MeImpG are sensitive to hydrolysis even in the solid state when stored in a vacuum container with desiccant, as was reported here. Given all this data and the presumed complexity of the prebiotic aqueous environment it seems unlikely that polymerisation of these precursors could compete with other nucleophiles in solution for substitution of the imidazolide moiety. The kinetically more stable di- and triphosphornucleotides, that could have been

readily synthesized from the imidazolide nucleotides and inorganic phosphate in prebiotic conditions, would require, unlike the 2-methylimidazole-activated nucleotides, some kind of catalyst for template directed polymerisation (probably "inorganic", like metal ions or mineral) but seem better candidates for the prebiotic job of synthesizing the first polyribonucleotides because of their kinetical stability to nucleophile substitution.

## Experimental

### Reagents

Nucleotides, KOD and NaOD (conc.) and reagents for 2-MeImpN synthesis were from Sigma; $D_2O$ (99.5% D), NaCl and buffers were from Merck.

### Synthesis of 2-MeImpA and 2-MeImpG

Preparation of the activated nucleotides was done by a procedure previously described.[2] The products were dried in a Petri dish for at least a week under vacuum with desiccant and stored also under these conditions in a small capped flask; yield was better than 95% in weight. However, the $^1$H NMR spectra revealed unaccounted singlets at 0.68 and 1.48 ppm (chemical shift reference: *tert*-butyl alcohol) that do not belong to the products. The intensity and chemical shifts of these signals didn't change with time. The $^{31}$P NMR spectra didn't show any unexpected signals. The chemical shifts observed for 2-MeImpA and 2-MeImpG spectra taken in the standard conditions for this work (see below) are presented in Fig. 1.

### NMR

Spectra were taken on a Bruker AC200 spectrometer in $D_2O$ solution using 5 mm NMR tubes. Chemical shift references were *tert*-butyl alcohol for $^1$H spectra (internal) and 65% phosphoric acid for $^{31}$P spectra (external). The $^1$H spectra were acquired at 200 MHz over 13 min with 15 ppm sweep width, 60° pulse width, 8 dummy scans and 160 scans. The $^{31}$P spectra were acquired at 81 MHz over 27 min with 37 ppm sweep width, 70° pulse width, 4 dummy scans and 512 scans.

Homonuclear ($^1$H–$^1$H and $^{31}$P–$^{31}$P) and heteronuclear ($^1$H–$^{31}$P) chemical shift correlation experiments were done after complete hydrolysis of the samples using H–H-COSY and H–X-COSY pulse sequences.[11]

### Sample preparation and standard experimental conditions for kinetic measurements

Buffer stock solutions were prepared by dissolving the required amounts of NaCl (0.5 M), $Na_2HPO_4/NaH_2PO_4$ (up to 0.5 M) and/or TRIS (up to 0.2 M) in $D_2O$. Buffer pD was adjusted with concentrated KOD and/or $DNO_3$ to 7.6 ($pD = pH_{measured} + 0.4$)[12] using a Crison Micro pH 2002 pHmeter and an Ingold U402 microelectrode calibrated with pH concentration buffers 4 and 7. Samples were prepared by dissolving in the NMR tube the required amount of 2-MeImpN in 0.5 ml of buffer to give a concentration of 0.020 ± 0.001 M. Spectra were taken with 2 to 3 h intervals at 308 K (35 °C) until 2-MeImpN $^1$H signals were no longer detected. Typically, 10 to 16 spectra were obtained; the time attributed to each spectrum was its acquisition finishing time. Mole fractions for each time were calculated based on the relative integrals of each nuclei signal set. Kinetic functions were fitted to the mole fraction data using linear and non-linear least-squares analysis.

### Molecular modelling

Semiempirical PM3 molecular modelling was done with PC Spartan Plus (Wavefunction, Inc.), ver. 1.4, using H–H repulsion correction, running in a 300 MHz Pentium II PC with Windows NT 4.0 Workstation OS.

## Acknowledgements

The authors would like to thank CEQUP and ICBAS for financial support and Professor Anake Kijjoa for fruitful discussions.

## References

1 T. Inoue and L. E. Orgel, *J. Mol. Biol.*, 1982, **162**, 201.
2 G. F. Joyce, T. Inoue and L. E. Orgel, *J. Mol. Biol.*, 1984, **176**, 279.
3 L. E. Orgel, *Nature*, 1992, **358**, 203.
4 D. Sievers and G. von Kiedrowski, *Nature*, 1994, **369**, 221.
5 T. Wu and L. E. Orgel, *J. Am. Chem. Soc.*, 1992, **114**, 317; 5496; 7963.
6 H. Sawai and M. Ohno, *Chem. Pharm. Bull.*, 1981, **29**, 2237; *Bull. Chem. Soc. Jpn.*, 1985, **58**, 361; H. Sawai, K. Higa and K. Kuroda, *J. Chem. Soc., Perkin Trans. 1*, 1992, 505.
7 J. P. Ferris and G. Ertem, *Origins Life Evol. Biosphere*, 1992, **22**, 369; K. Kawamura and J. P. Ferris, *J. Am. Chem. Soc.*, 1994, **116**, 7564; P. Ding, K. Kawamura and J. P. Ferris, *Origins Life Evol. Biosphere*, 1996, **26**, 151.
8 A. Kanavarioti, C. F. Bernasconi, D. L. Doodokyan and D. S. Alberas, *J. Am. Chem. Soc.*, 1989, **111**, 7247.
9 A. Kanavarioti, C. F. Bernasconi, D. L. Doodokyan and E. E. Baird, *J. Am. Chem. Soc.*, 1993, **115**, 8537.
10 J. Ferris, *Nature*, 1994, **369**, 184.
11 W. P. Awe, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.*, 1976, **64**, 2229; A. Bax and G. Morris, *J. Magn. Reson.*, 1981, **42**, 501.
12 P. K. Glasoe and F. A. Long, *J. Phys. Chem.*, 1960, **64**, 188.

*Paper 9/00884E*