Supporting Information

Synthesis of Imidazole-Activated Ribonucleotides Using Cyanogen Chloride

Ruiqin Yi, Yayoi Hongo, Albert C. Fahrenbach[*]

CONTENTS

1.   General Methods

1a. General considerations. All reactions were carried out in 18 MΩ water using a MilliQ (Merck Millipore, Billerica, MA, USA) purification system. All reagents were purchased from Sigma Aldrich (Tokyo, Japan) or TCI (Tokyo, Japan) and used without further purification. The concentrations of aqueous solutions of imidazole were calibrated using their UV absorption at 205 nm using a molar extinction coefficient[1] of 4728 M^{-1} cm^{-1}. The concentrations of aqueous NaOCl solutions were calibrated by their absorbance at 293 nm using a molar extinction coefficient[2] of 365 M^{-1} cm^{-1}. UV absorption spectra were measured on a Thermo Scientific Nanodrop 2000c Spectrophotometer (Waltham, MA USA). Automated syringe pumps were purchased from YMC (Kyoto, Japan). Flash chromatography was carried out with a Biotage SP1 Flash Purification system (Charlottesville, VA USA) equipped with a Snap Ultra C_{18} 120 g column.

1b. NMR spectroscopy. NMR spectra were acquired on a Bruker (Billerica, MA, USA) Avance[III] 400 MHz NMR spectrometer equipped with a 5 mm z-gradient PABBO BB-1H/D Z-GRD probe (400 MHz for ^{1}H, 100 MHz for ^{13}C, 160 MHz for ^{31}P). Proton, phosphorus and carbon chemical shifts are reported in parts per million (ppm) values on the δ scale. ^{1}H NMR spectra were internally referenced to either the residual DHO signal (δ = 4.79 ppm) or the d_6-DMSO signal (δ = 2.50 ppm). ^{31}P NMR spectra were referenced to the phosphorus resonances of the CMP, UMP, AMP, GMP or 2′,3′-di-*O*-acetyl-cytidine-5′-monophosphate starting material, which were all set to 0 ppm. All NMR spectra were recorded at room temperature and were analyzed using MestReNova (MestreLab Research, Santiago de Compostela, Spain). Baseline subtraction was carried out before all resonance integrations, which were measured by carrying out a peak fitting analysis.

1c. Gas chromatography mass spectrometry. Volatiles in the reactions of NaCN with NaOCl were analyzed with a Shimadzu (Kyoto, Japan) GCMS-QP2010 Ultra gas chromatrograph-mass

spectrometer (GCMS). Head-space sampling of the reaction vials was carried out using a PAL (prep and load solution) RTC autosampler (CTC Analytics, Zwingen, Switzerland). Data processing and analysis was carried out using the LabSolution_GCMSsolution version 4.11 software package (Shimadzu). The injection method used an inlet temperature of 60 ^{o}C and a source temperature of 200 ^{o}C. The scan range was set to m/z 10–250 Da. Chromatography was performed on an Rtx-WAX® column (60 m, 0.3 mm I.D., 0.5 mm df). Helium was used as the carrier gas at a flow rate of 40 cm sec^{-1}. The oven temperature program for heating the column was as follows: 2 min hold at 60 ^{o}C, then from 60 to 100 ^{o}C at 10 ^{o}C min^{-1}, then from 100 to 230 ^{o}C at 200 ^{o}C min^{-1}, and finally the temperature was maintained at 230 ^{o}C for 2 min. Compounds were identified by comparison of their electron impact fragmentation spectra with standards from the NIST standard reference library.

1d. Liquid chromatography mass spectrometry (LCMS). High-resolution mass data analyses were performed on a Xevo-Q-TOF-MS platform (Waters Corporation, Manchester, UK) combined with a ultra-performance liquid chromatography (UPLC) system (Acquity H-class; Waters Corporation, MA, USA). Data analysis was carried out using MestReNova. Liquid chromatography was carried out on a CORTECS UPLC C_{18} column (1.6 um, 2.1 mm x 150 mm, Waters) at a flow rate of 0.2 mL min^{-1}. Solvent A was 0.1% trifluoroacetic acid in water and solvent B was acetonitrile. Solvent A was maintained at 100% for two minutes, then solvent B was ramped from 0 to 30% from 2.0 to 5.0 minutes, and then to 80% from 5.0 to 6.0 minutes. Solvent B remained at 80% for one minute and then ramped back down to 0% over one minute. Finally, 100% solvent A was eluted for an additional two minutes. The column was heated to 35 ^{o}C, and all samples were kept at 4 ^{o}C while queued in the autosampler. The injection volumes were either

1.0 or 4.0 uL. Under positive mode ESI conditions, a voltage of 3.0 kV was applied to the stainless-steel electrospray ionizer. The TOF analyzer was set to sensitivity mode with a resolving power of 22,000, and the set m/z range of 20–1500 was calibrated with sodium formate. The desolvation gas (nitrogen) was used at a flow rate of 500 L h^{-1}, and the source and desolvation temperatures were set to 100 °C and 250 °C, respectively.

1e. Synthesis. Authentic standards of diimidazole imine (Im$_2$CNH) were synthesized according to previously reported procedures[3] by refluxing 1 molar equivalent of cyanogen bromide with 3 molar equivalents of imidazole in dichloromethane for 30 minutes. Two rounds of recrystallization from cold dichloromethane (–20 °C) were found to be necessary in order to obtain the target product in >95% purity. Standards for N-cyanoimidazole were synthesized from a previously reported procedure[4] by reaction of 1 molar equivalent of cyanogen bromide with 1 molar equivalent of imidazole in acetonitrile at room temperature. Standards for the four ribonucleoside 5′-phosphorimidazolides (ImpC, ImpU, ImpA, ImpG) were synthesized according to a previously reported procedure[5] that relied on 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC). For each reaction, an aqueous solution of 0.1 M CMP, UMP, AMP or GMP, respectively, with 1 M imidazole and 1 M EDC was prepared and adjusted to pH 6 without hesitation using concentrated HCl. The solutions were allowed to react for a few minutes at room temperature, at which point ~95% conversion to the respective imidazolides had taken place. These freshly prepared solutions were then used without further purification as standards for NMR spectroscopy.
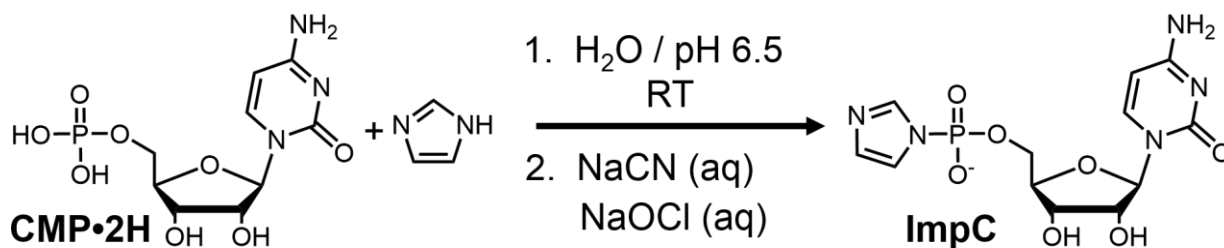
Synthesis of ImpC from Im, NaCN and NaOCl: A 5 mL aqueous solution containing 0.1 M CMP•2H (162 mg, 0.50 mmole) and 2 M imidazole (680 mg, 10 mmole) was prepared and adjusted to pH 6.5 using concentrated HCl, and then added to a round bottom flask fitted with a magnetic stir bar. Separately, a 5 mL aqueous solution of 1 M NaCN (245 mg, 5 mmole) was freshly prepared, and 5 mL of an aqueous solution of 1 M NaOCl was obtained by the appropriate dilution of a calibrated stock solution. The NaCN and NaOCl solutions were loaded into separate all-plastic syringes, and the tips of the syringes were fitted with PTFE tubings that led *directly* into the solution containing CMP•2H and imidazole placed in the round-bottom flask. (See Figure S15 for the results of a reaction in which the PTFE tubings are placed *above* the reaction mixture). The solutions of NaCN and NaOCl were then added via automated syringe pumps at a rate of 5 mL h^{-1} while the solution was mildly stirred at room temperature. Small samples of the reaction mixture were taken at regular intervals and analyzed by NMR spectroscopy by adding enough D$_2$O to make a ~10% solution. Addition of freshly prepared authentic standards of ImpC added directly to the NMR tube confirmed the production of ImpC. Samples were also analyzed by high-resolution (Q-TOF) LCMS, and these experiments also confirmed the synthesis of the target product. *CAUTION: NaCN is highly toxic as is the volatile ClCN produced in situ. The reaction must be performed within a well ventilated fume hood.*
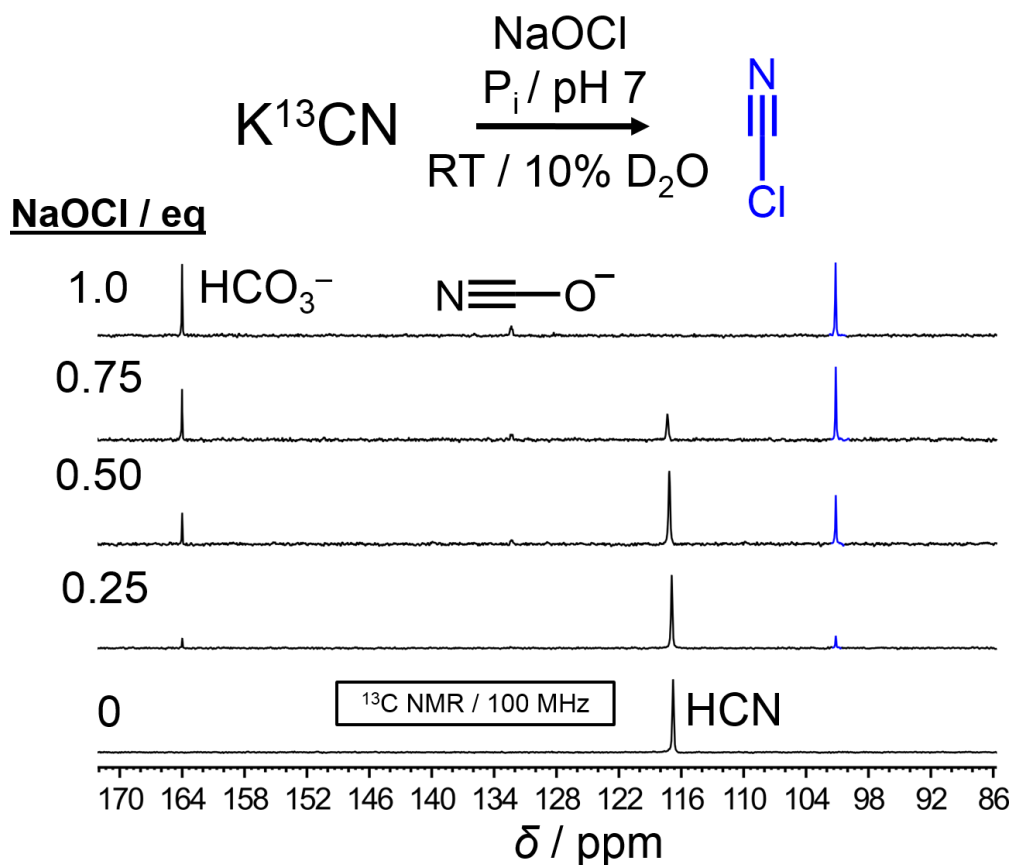
2. Figures

Figure S1. Titration of NaOCl into a solution of $K^{13}CN$ at pH 7 monitored by ^{13}C NMR spectroscopy yields ClCN. The titration was carried out in 10% D_2O at room temperature starting with a solution of 0.1 M $K^{13}CN$ and 0.5 M sodium phosphate buffered at pH 7. Additions of NaOCl were carried out using a concentrated stock solution of NaOCl also in 10% D_2O that was calibrated based on UV absorbance measurements. The resonance colored in blue is assigned to ClCN, the chemical shift of which is consistent[6] with previous reports. Resonance assignments for bicarbonate and cyanate anions were carried out by additions of authentic standards. Cyanate is likely the hydrolysis product[7] of ClCN perhaps catalyzed by the phosphate buffer, and bicarbonate is the hydrolysis product[8] of cyanate. P_i: inorganic phosphate.
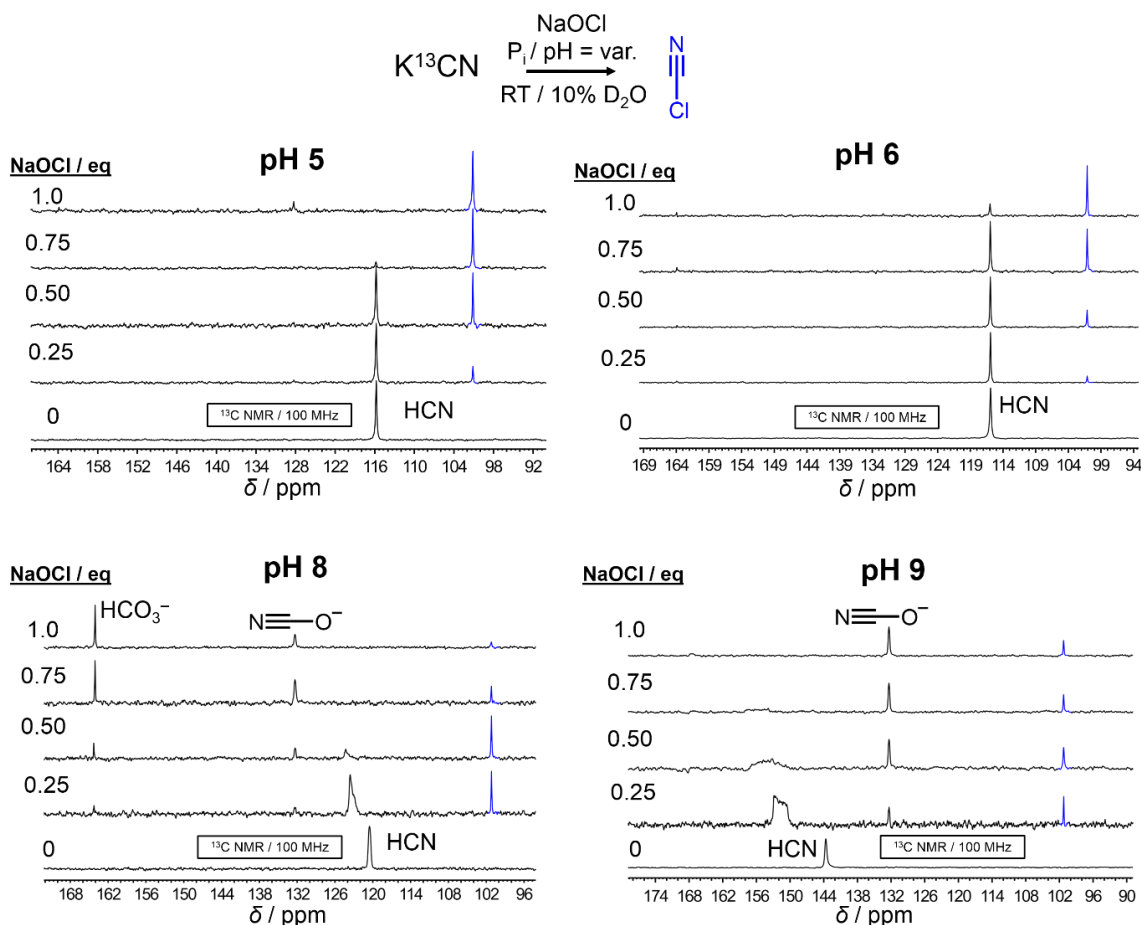
Figure S2. Titration of NaOCl into a solution of $K^{13}CN$ monitored by ^{13}C NMR spectroscopy at different pH values. The titrations were all carried out in 10% D_2O at room temperature starting with a solution of 0.1 M $K^{13}CN$ and 0.5 M sodium phosphate buffered at the pH values shown in each panel. Additions of NaOCl were carried out using a concentrated stock solution of NaOCl also in 10% D_2O that was calibrated based on UV absorbance measurements. Resonances colored in blue are assigned to cyanogen chloride. At pH 9, the hydrolysis of cyanate to bicarbonate is extremely slow and not observed on the timescale of these experiments. At pH 5 and 6, hydrolysis of ClCN is also relatively slow. For each titration, the acquisition (32–64 scans) of the successive titrative spectra began on average 3–5 minutes in real time after the start of the acquisition of the spectrum proceeding it.
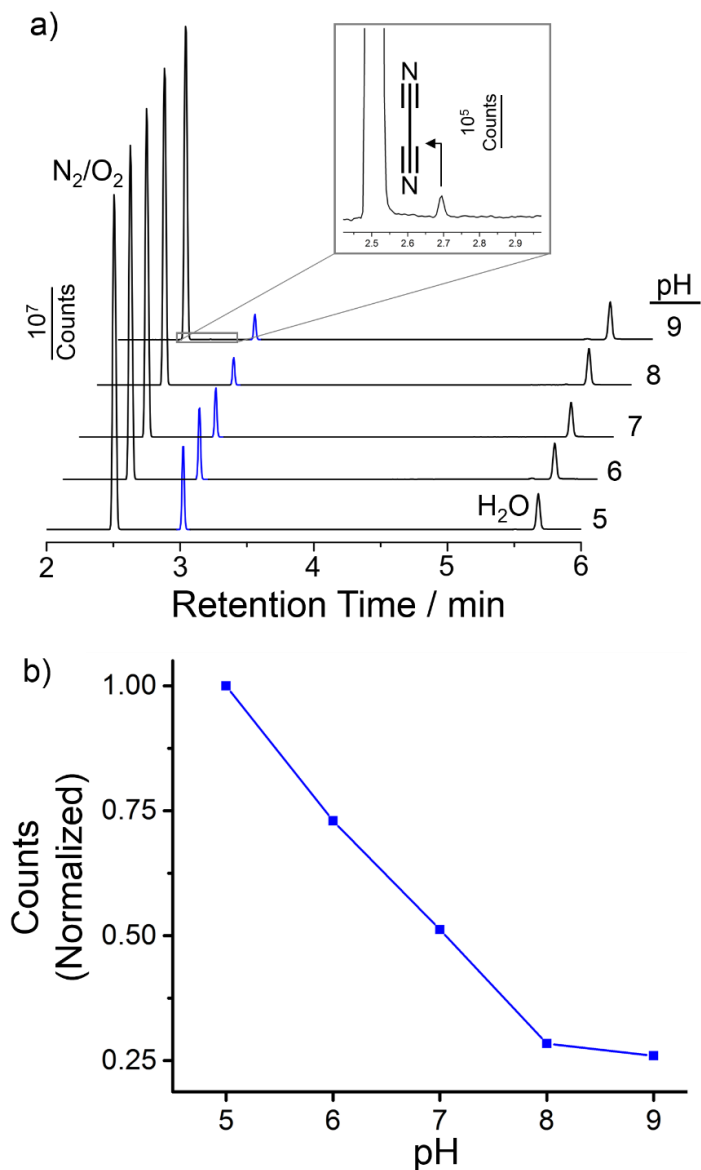
S7

Figure S3. Monitoring the formation of ClCN at different pH values using GCMS. Solutions containing 1 mL of 25 mM NaCN and 100 mM of sodium phosphate buffered at pH 5, 6, 7, 8, or 9 were prepared inside of GCMS vials intended for head-space sampling. The vials were capped and then ~1.5 eq of NaOCl with respect to NaCN was added using a concentrated stock solution. The vials were shaken for one minute by hand, and then the head-space of the vials were analyzed by GCMS. a) Total ion chromatograms from 2 to 6 minutes retention time recorded on the head-space of the reactions described above at the different pH values shown. The inset shows a small peak with a mass spectrum that matches cyanogen gas $(CN)_2$ observable at pH 9. b) Normalized counts relative to pH 5 for the peak assigned to cyanogen chloride at different pH values. The yield of ClCN appears to decrease at higher pH likely as a result of hydrolysis to cyanate.
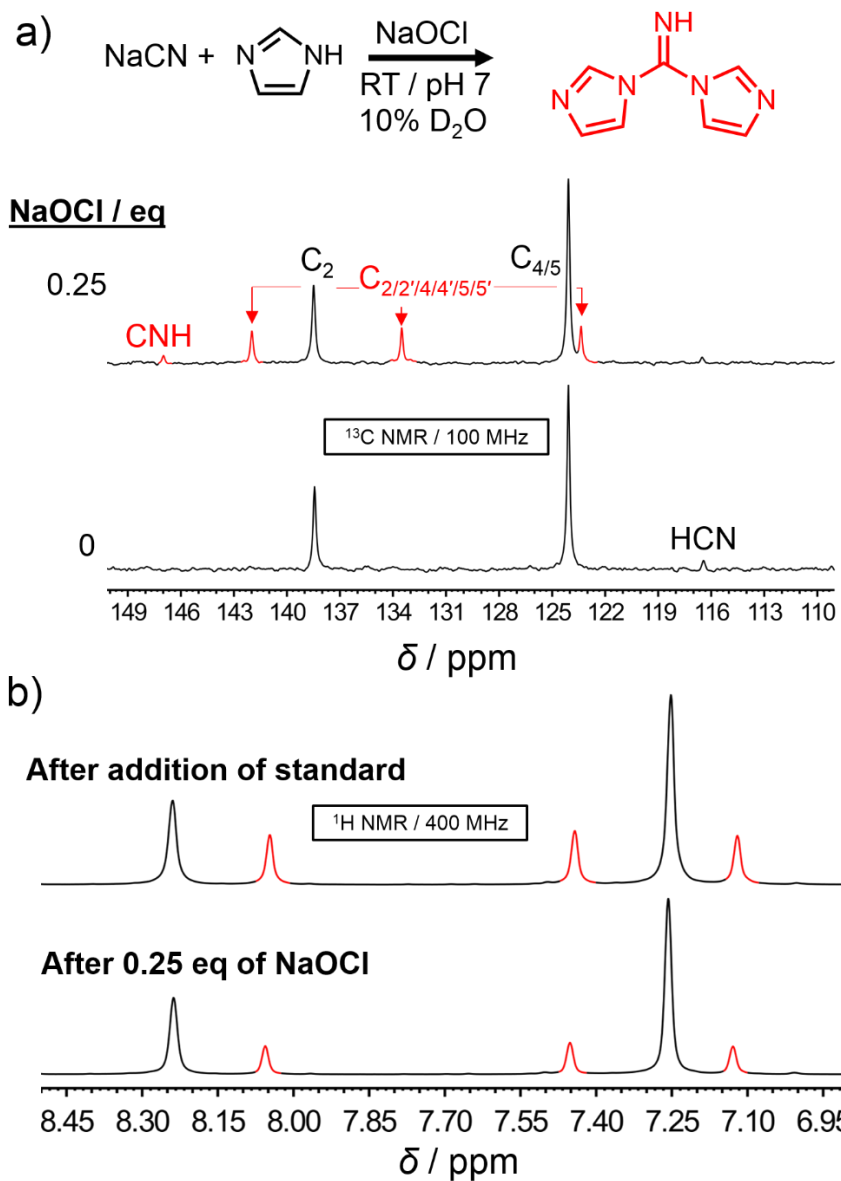
Figure S4. Carbon and proton NMR spectroscopies confirm the synthesis of Im_2CNH. a) The ^{13}C NMR spectra of a 10% D_2O solution at pH 7 containing 1 M of Im and 0.4 M of NaCN before (bottom trace) and after titrating up to 0.25 eq of NaOCl (top trace). b) An authentic standard of Im_2CNH was then added to the same solution shown in a) after adding 0.25 eq of NaOCl. The resonances colored in red increased providing additional evidence of their assignments.
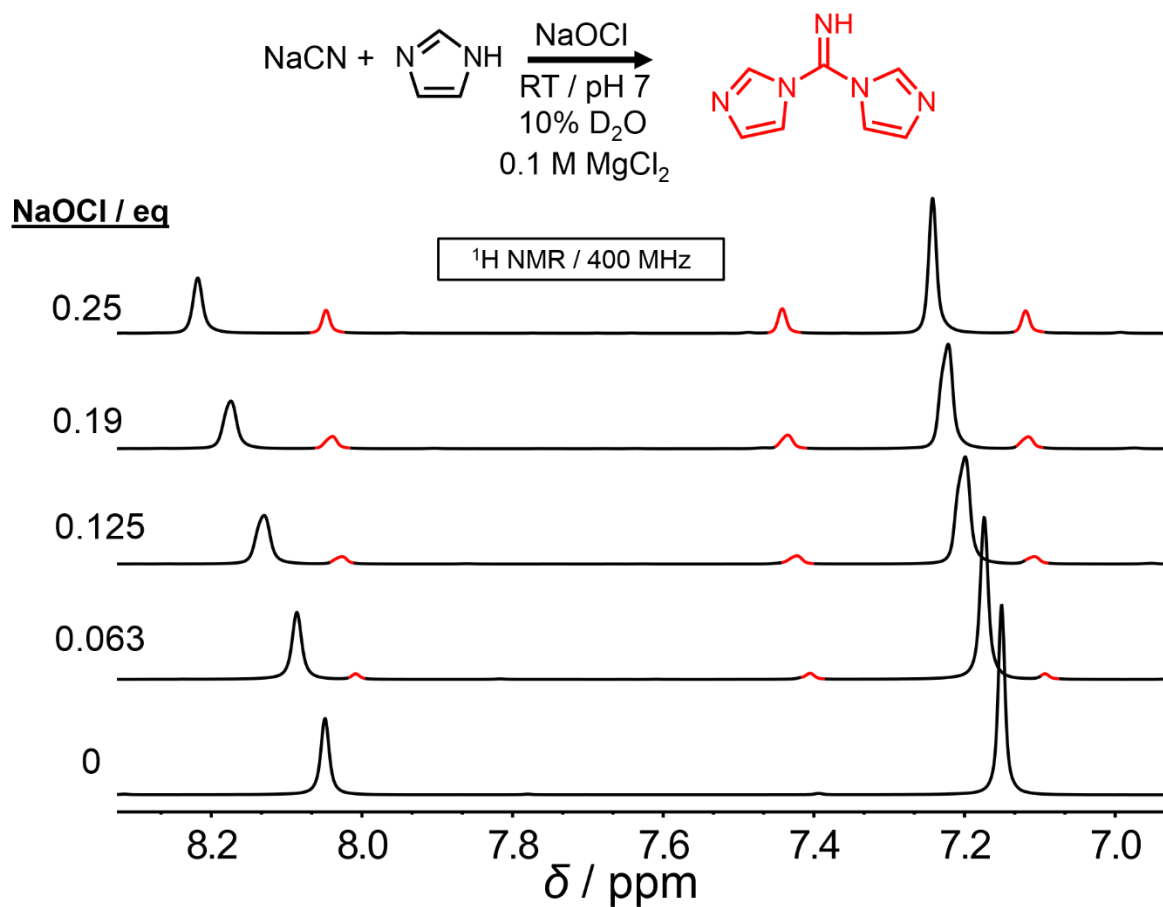
Figure S5. Proton NMR spectroscopy confirms the synthesis of Im_2CNH is not affected by Mg^{2+}. The ^1H NMR spectra of a 10% D_2O solution at pH 7 containing 1 M of Im, 0.4 M of NaCN and 0.1 M of $MgCl_2$ were recorded during a titration up to 0.25 eq of NaOCl. After the addition of 0.25 eq of NaOCl, the yield of Im_2CNH was determined by relative resonance integration to be ~60% based on NaOCl as the limiting reagent.
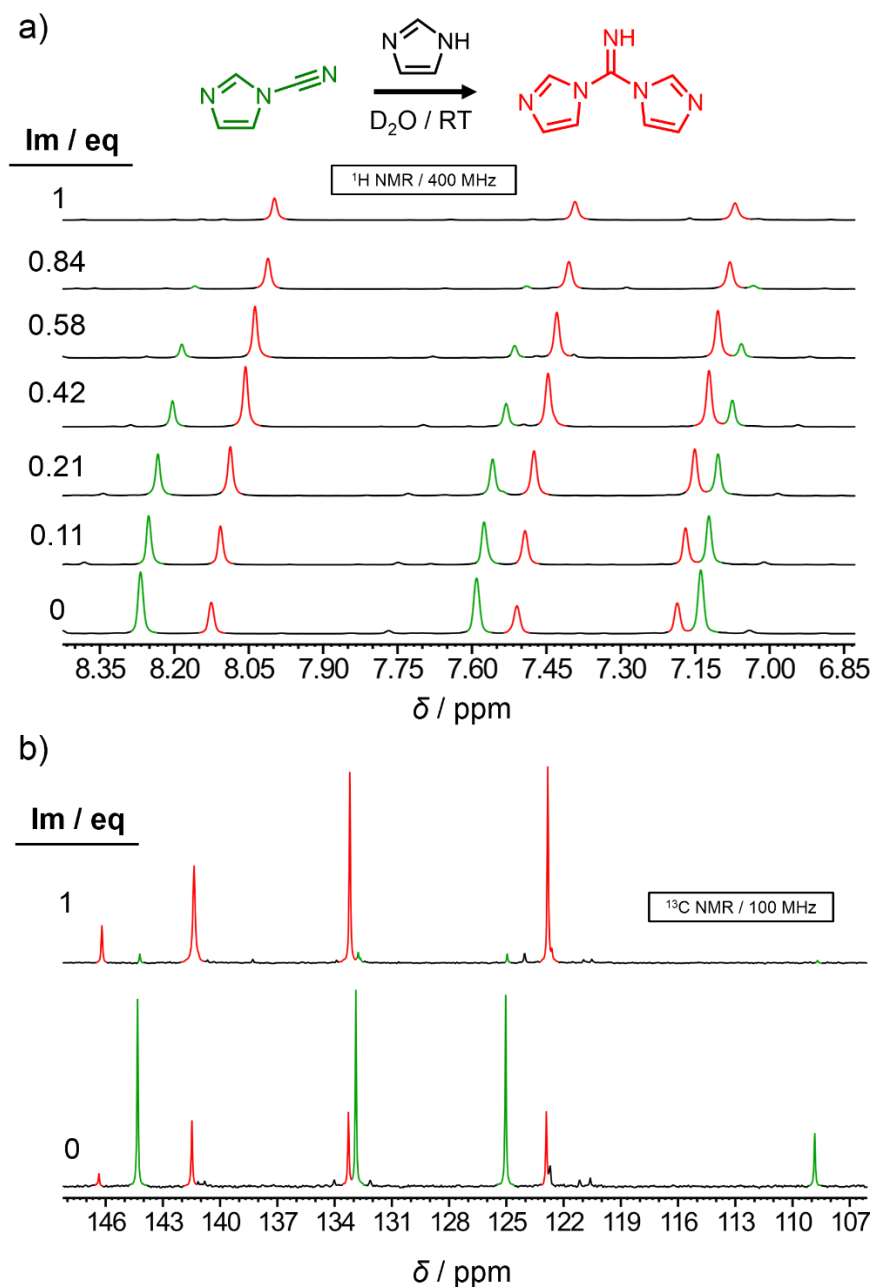
Figure S6. Addition of imidazole to ImCN results in the formation of Im_2CNH. a) A freshly prepared sample of ImCN was synthesized according to Blackman et al.[4], and then was dissolved in 450 microL of D_2O making a ~1 M solution. The ^1H NMR spectra of the sample obtained reveals a mixture of ImCN and Im_2CNH. A concentrated solution of Im was then titrated into the solution as the ^1H NMR spectra were recorded. b) The ^{13}C NMR spectra of the same solution in a) before and after addition of 1 eq of Im.
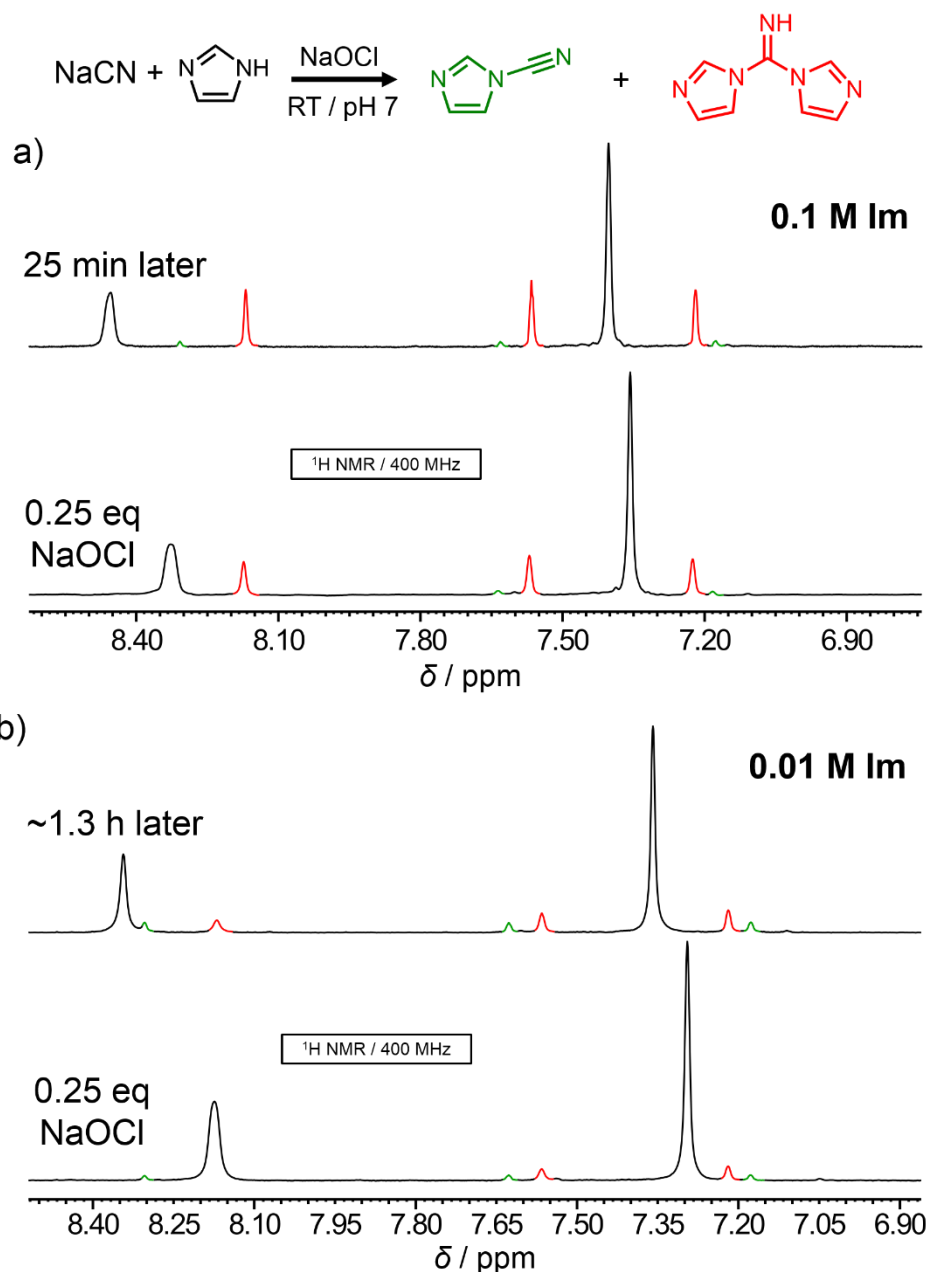
Figure S7. Titrations involving lower initial concentrations of Im result in a greater proportion of ImCN, but similar total yields of ImCN and Im_2CNH are obtained after longer reaction times. a) The ^1H NMR spectrum of a 10% D_2O solution at pH 7 containing 0.1 M of Im and 0.1 M of NaCN after a titration of NaOCl up to 0.25 eq, and the spectrum obtained of the same solution 25 min later. The yields of Im_2CNH were 51% and 63%, respectively. b) The ^1H NMR spectrum of a D_2O solution at pD 7 containing 0.01 M of Im and 0.01 M of NaCN after a titration of NaOCl up to 0.25 eq, and spectrum obtained of the same solution ~1.3 h later. A significant proportion of ImCN was observed. The yields of Im_2CNH were 28% and 33%, respectively, while the yields for ImCN were 19% and 33%, respectively. All yields were determined by relative resonance integrations based on NaOCl as the limiting reagent.
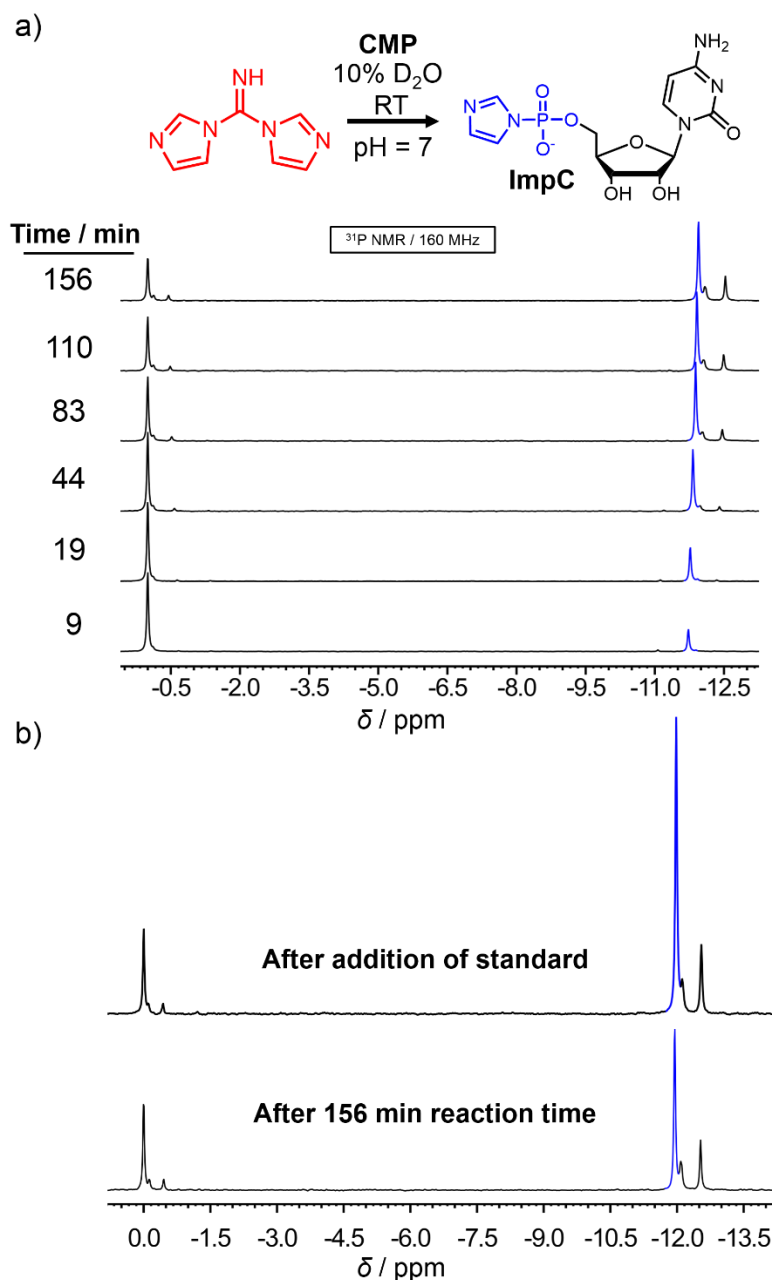
Figure S8. Im$_2$CNH is an activating agent for CMP. a) A 500 uL solution of 0.1 M CMP in 10% D$_2$O at pH 7 was prepared, and then 81 mg (10 eq) of Im$_2$CNH was added as a solid, and the pH of the solution was adjusted back to 7 without hesitation. The ^{31}P NMR spectra of the solution were then recorded over time. b) After 156 min of reacting, a synthetically prepared standard of ImpC was added to the reaction mixture shown in a). The resonance colored in blue increased in intensity after the addition confirming that it arises from the phosphorous atom of ImpC.
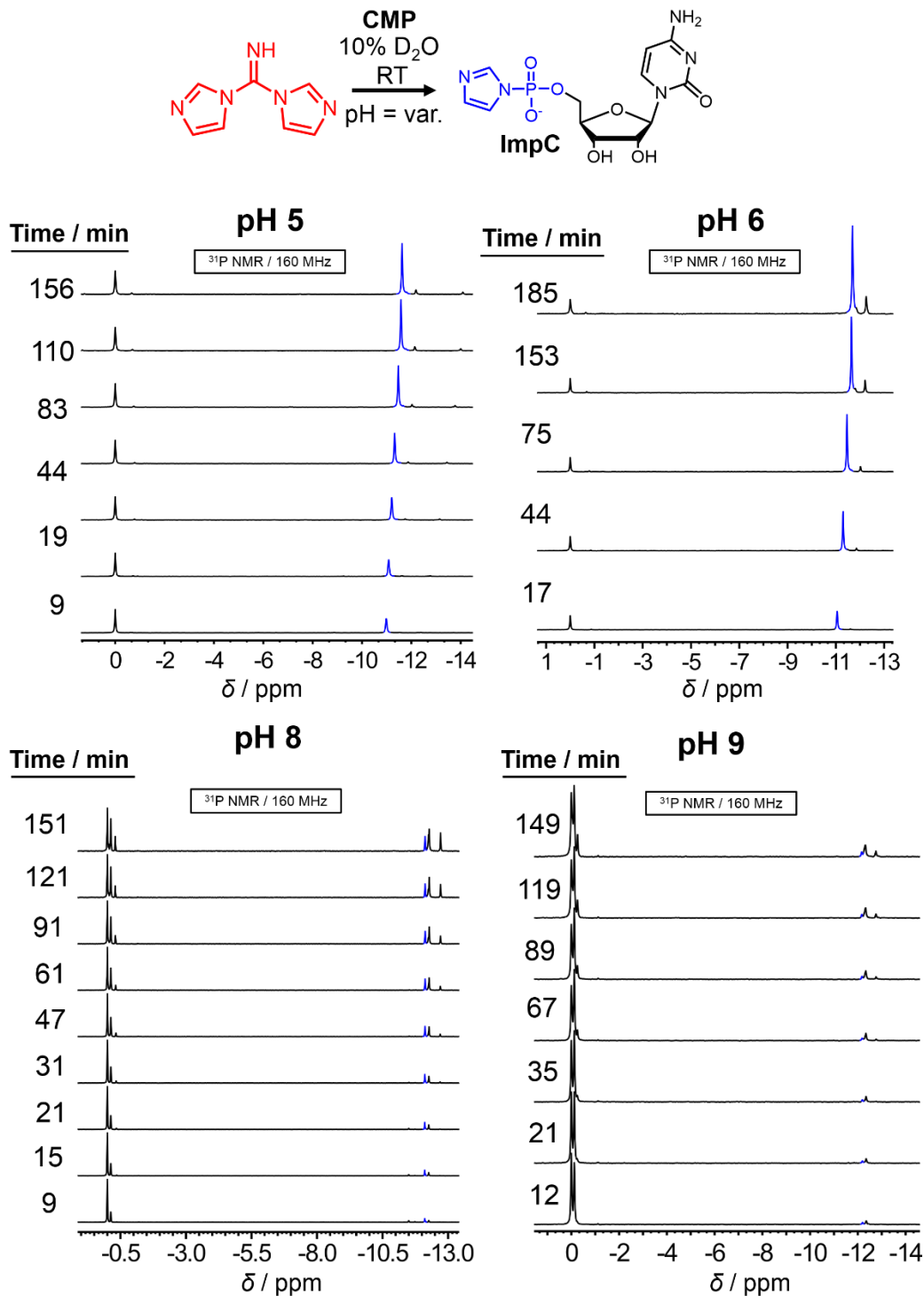
Figure S9. The synthesis of ImpC by Im₂CNH at different pH values. All reactions were performed in the same manner as described in Figure S8 except at the pH values shown above in each panel. Identification of the resonance arising from ImpC was carried out by addition of authentic standard for each pH tested. These data were used to generate Figure 2a in the main text.
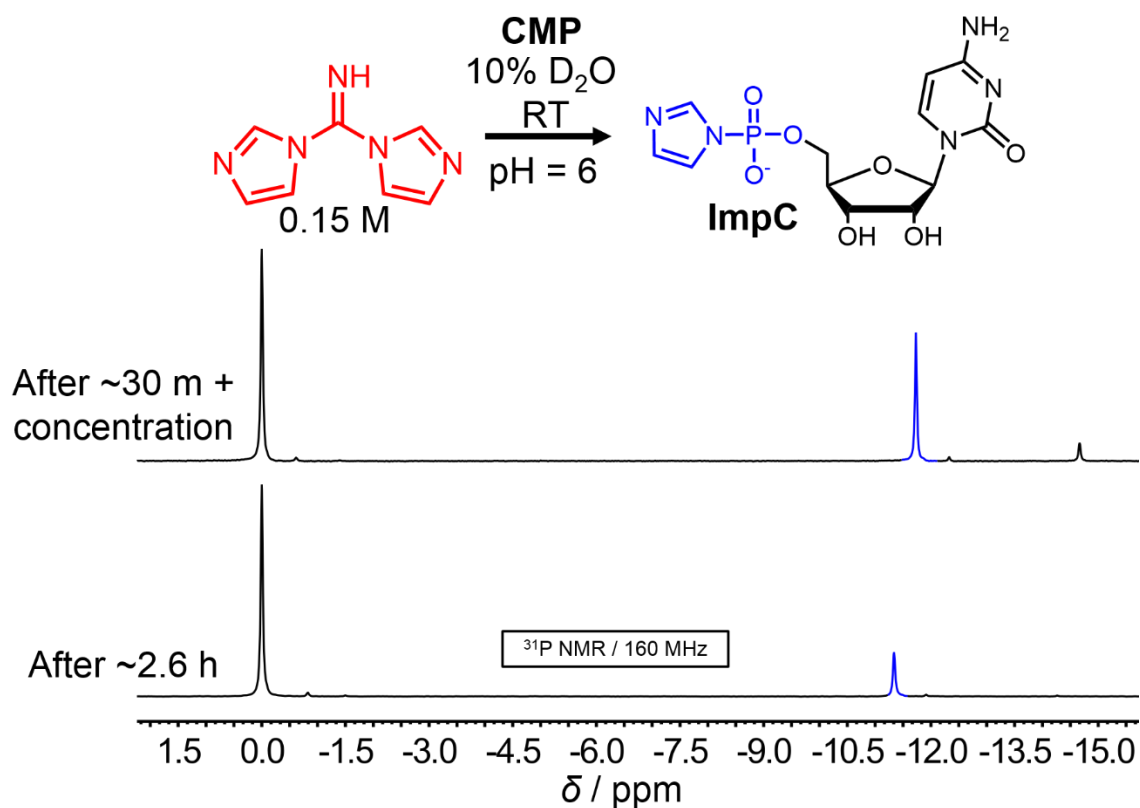
Figure S10. A lower concentration of Im$_2$CNH results in a decreased yield of ImpC. A solution of 0.1 M CMP and 0.15 M Im$_2$CNH was prepared in 10% D$_2$O at pH 6, and the reaction was monitored using ^{31}P NMR spectroscopy over time. After 2.6 h of reaction time, analysis of the ^{31}P NMR spectrum reveals that the yield of ImpC is 18% (bottom trace). Concentration of the mixture under vacuum tends to increase the yield. A reaction prepared in a similar fashion was reacted for 30 minutes, then concentrated under vacuum at 35 °C for about 15 min. The residual solid was dissolved in D$_2$O and the ^{31}P NMR spectrum recorded. Analysis reveals the yield of ImpC is about 33% (top trace). We suspect that the lower yields of ImpC in comparison to when 1 M of Im$_2$CNH is used is a result of the lower rate of ImpC formation, which must compete with the constant background hydrolysis of ImpC back to CMP.
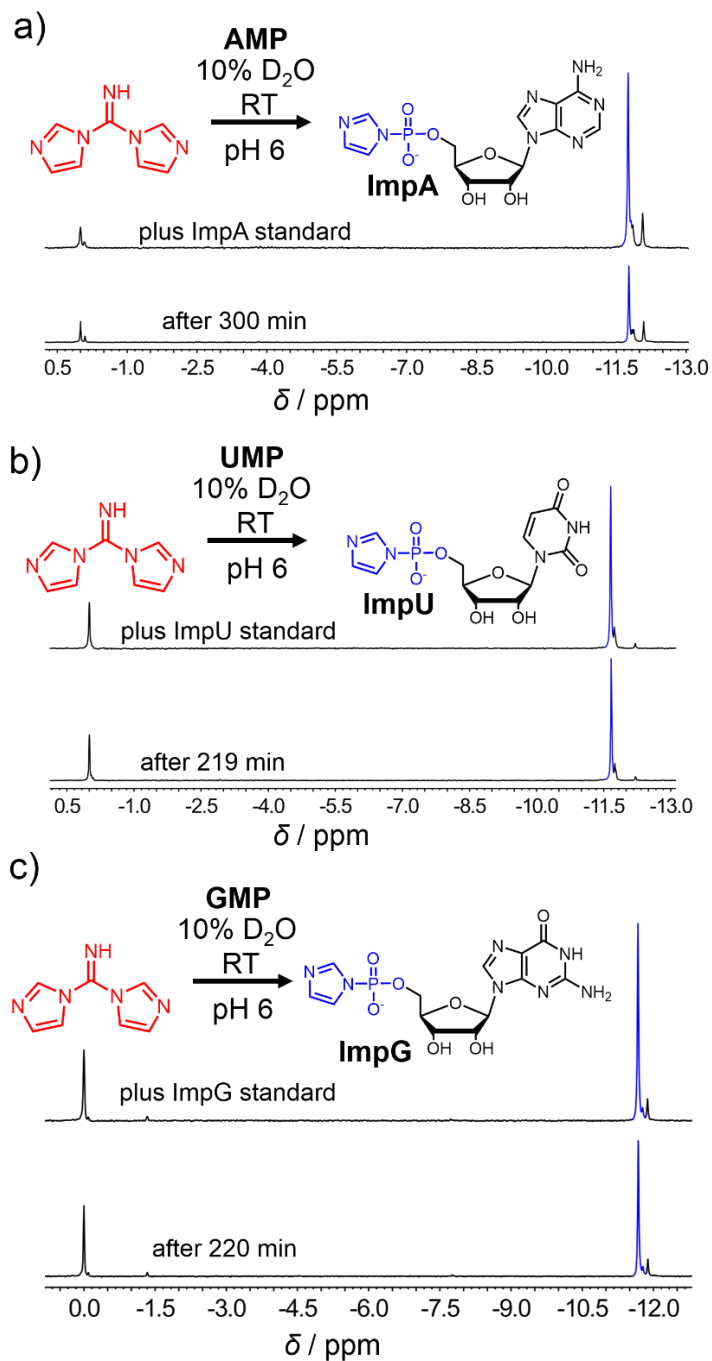
Figure S11. Im₂CNH enables activation of all four canonical RNA monomers. Reactions of 1 M Im₂CNH at pH 6 in 10% D₂O at RT with 0.1 M of a) AMP, b) UMP and c) GMP. The reactions were monitored by ^{31}P NMR spectroscopy. The reactions were allowed to proceed for the times shown, after which additions of authentic standards of ImpA, ImpU and ImpG, respectively, led to observed increases in the relative intensities of the resonances colored in blue.
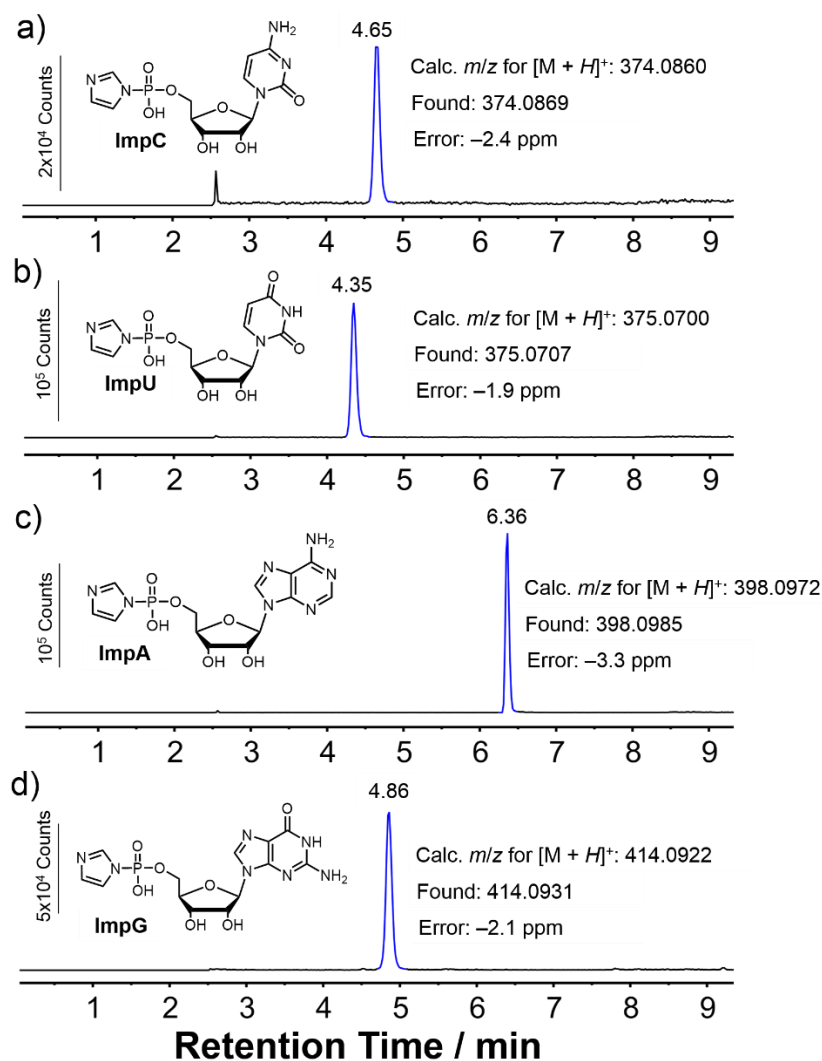
Figure S12. LCMS analysis provides evidence for ribonucleoside 5′-phosphorimidazolide synthesis from Im, NaCN and NaOCl. LCMS analyses were carried out on 1000-fold diluted samples of reaction mixtures carried out using syringe pumps in a similar manner as described in the General Methods section after all the NaCN and NaOCl had been added. Briefly, each reaction was carried out by preparing a 2.5 mL solution containing one of the ribonucleoside 5′-monophosphates (0.1 M) and imidazole (2 M), and was adjusted between pH 6 and 6.3 with concentrated HCl. To these solutions were pumped in separate 2.5 mL solutions of NaCN (1 M) and NaOCl (1 M) at 2.5 mL h^{-1}. Extracted ion chromatograms for the $[M + H]^+$ calculated masses of a) ImpC taken at 115 min reaction time, b) ImpU taken at 75 min reaction time, c) ImpA taken at 90 min reaction time and d) ImpG taken at 80 min reaction time. All chromatograms were extracted with a tolerance of ±0.1 Da. For each RNA monomer, analysis of reaction samples that were combined with authentic standards prepared as described in the General Methods section yielded extracted ion chromatograms displaying only a single peak corresponding to the $[M + H]^+$ m/z values of the target imidazolides.
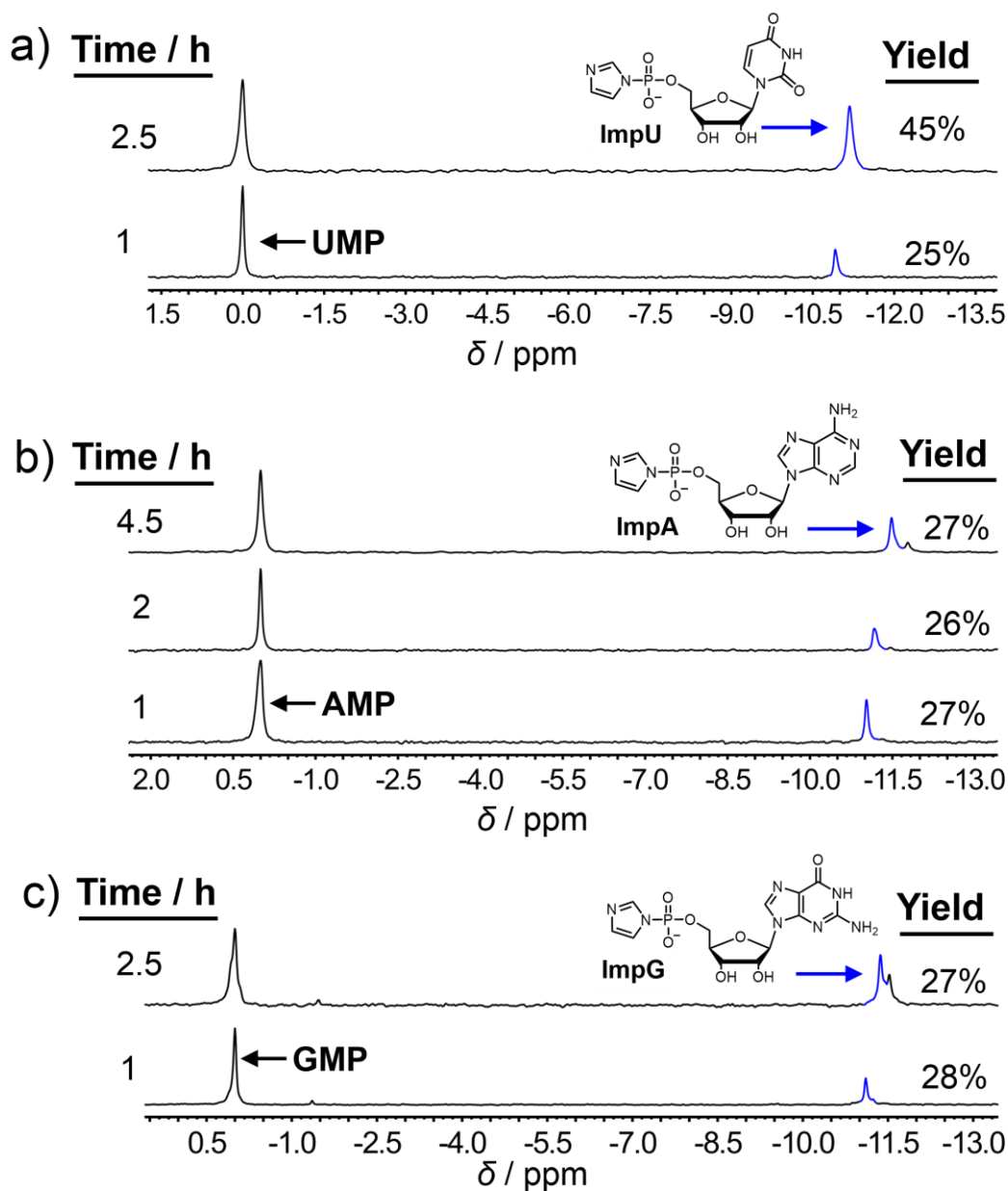
Figure S13. ^{31}P NMR spectroscopy confirms the synthesis of ribonucleoside 5'-phosphorimidazolides. ^{31}P NMR analyses were carried out on 400 *u*L samples of reaction mixtures that were taken at the times shown and combined with 50 *u*L of D_2O. All reactions were carried out using syringe pumps in the same manner as described in the General Methods section. Briefly, each reaction was carried out by preparing a 2.5 mL solution containing one of the ribonucleoside 5'-monophosphates shown (0.1 M) and imidazole (2 M), and was adjusted to pH 6 with concentrated HCl. To these solutions were pumped in separate 2.5 mL solutions of NaCN (1 M) and NaOCl (1 M) at 2.5 mL h^{-1}. After 1 hour, all the NaCN and NaOCl solutions had been added. Partial ^{31}P NMR traces recorded at room temperature for reactions containing a) UMP, b) AMP, and c) GMP. Assignment of all resonances colored in blue was carried out by addition of authentic standards of the respective phosphorimidazolides.
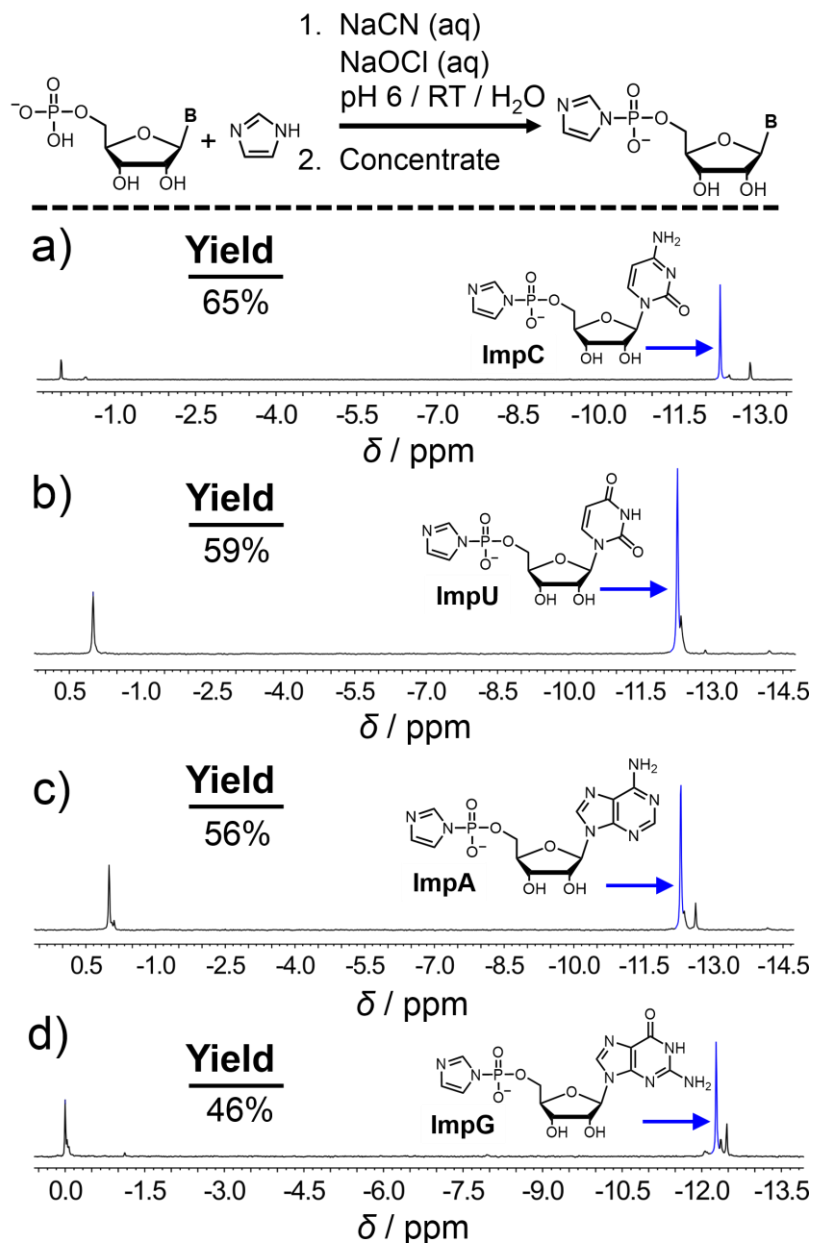
Figure S14. Concentrating the reaction mixtures tends to increase the yield of the ribonucleoside 5′-phosphorimidazolides. Reactions were carried out in the same way as described in Figure S13. After one hour, when all the NaCN and NaOCl had been added, 2.5 mL of the reaction mixture was transferred to a round bottom flask and concentrated under vacuum at 40 °C for 15 minutes, which was long enough to bring the mixture to dryness. The residual solids were then redissolved in 500 uL of D_2O and their ^{31}P NMR spectra recorded. Partial ^{31}P NMR spectra for reactions containing either a) CMP, b) UMP, c) AMP or d) GMP as the starting material. The resonances colored in blue were identified by addition of authentic standards.
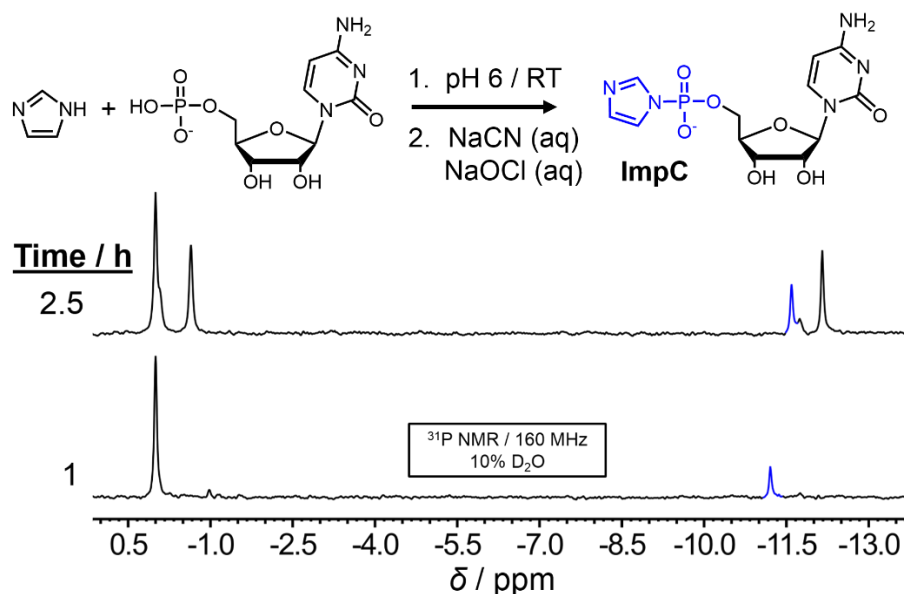
Figure S15. When the PTFE tubings are placed above the reaction mixture and the solutions of 1 M NaCN and NaOCl are allowed to drip in, the yield of side-products increase. A similar reaction as described in Section 1e was carried out, but with the PTFE tubings placed above the reaction mixture containing Im and CMP. After 1 hour, when all the NaCN and NaOCl solutions had been added, a 450 uL sample of the reaction mixture was combined with 50 uL of D_2O and the ^{31}P NMR spectra were recorded over time. After 2.5 h of total reaction time, a significantly greater amount of side-products were observed to form. We hypothesize this is a result of the formation of a relatively large amount of dissolved cyanogen $(CN)_2$. The volatility of either ClCN or HOCl/Cl_2 leads to significant dissolution within the hanging droplet of 1 M NaCN, at which point the rapid formation of cyanogen is thought to take place. Cyanogen then reacts with the 2′- and 3′-hydroxyls of CMP and ImpC. When a similar reaction was performed on the protected 2′,3′-di-*O*-acetyl-cytidine-5′-monophosphate, no side products were observed. See Figure S16 for more details.
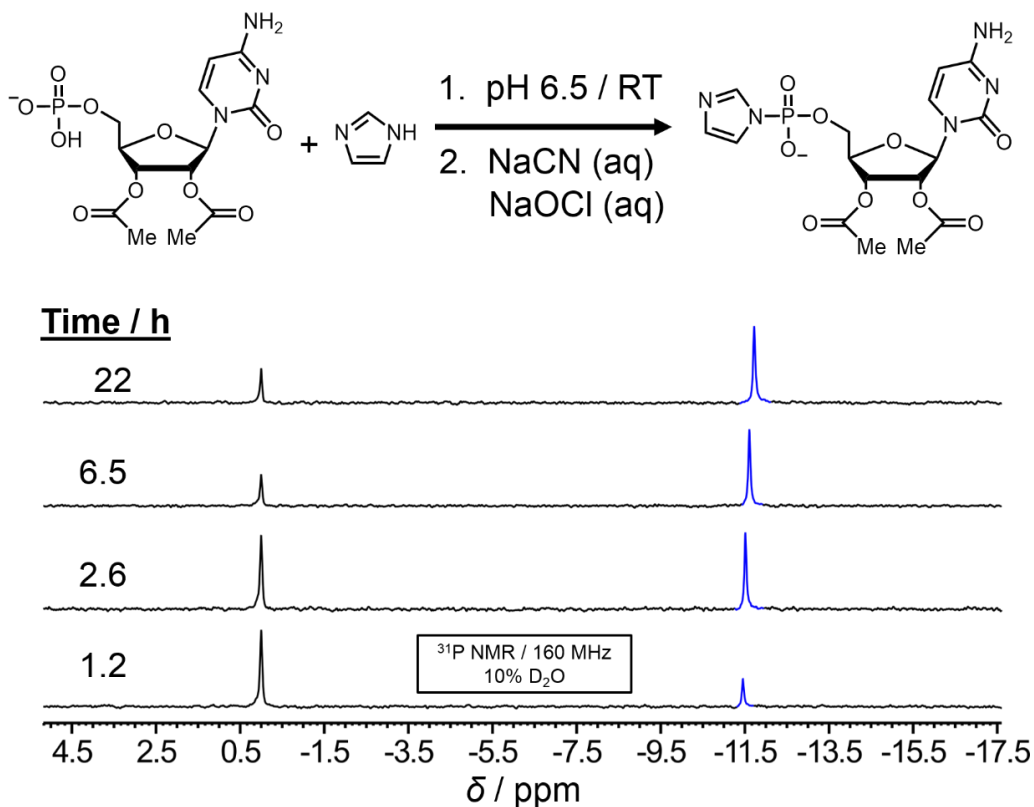
Figure S16. No side products are observed by ^{31}P NMR spectroscopy if 2′,3′-di-*O*-acetyl-cytidine-5′-monophosphate is employed in the pump reaction with NaCN and NaOCl. The 2′,3′-di-*O*-acetyl-cytidine-5′-monophosphate[9] was synthesized and purified according to a previously reported procedure. Then, a 1 mL solution containing 0.05 M of this protected CMP with 2 M Im was prepared at pH 6.5. To this mixture was pumped in separate 1 mL solutions of 1 M NaCN and 1 M NaOCl at 1 mL h^{-1} in a similar fashion as described in Section 1e. After one hour, when all the NaCN and NaOCl had been added, a 400 uL sample of the reaction mixture was combined with 50 uL D$_2$O, and the ^{31}P NMR spectra shown above were recorded over time. After 22 h, analysis of the ^{31}P NMR spectrum revealed a yield of 72%, and no other resonances arising from side-products were observed. The reaction mixture was purified by flash chromatography on a C$_{18}$ column using a binary solvent elution gradient (solvent A: H$_2$O; solvent B: MeCN). Fractions containing the target compound were combined, and 500 uL of *d$_6$*-DMSO was added. Then, the MeCN and majority of the H$_2$O were removed under vacuum at 35 °C, after which time the ^{1}H and ^{13}C NMR spectra, as well as the high-resolution mass spectrum, were recorded without hesitation. ^{1}H NMR (400 MHz, *d$_6$*-DMSO): δ = 7.92 (s, 1H), 7.53 (d, *J* = 7.5 Hz, 1H), 7.20 (s, 1H), 7.02 (s, 1H), (d, *J* = 5.6 Hz, 1H), 5.87 (d, *J* = 3.6 Hz, 1H), 5.17 (t, *J* = 5.6 Hz, 1H), 5.08–5.03 (m, 1H), 4.15 (dq, *J* = 5.2, 2.6 Hz, 1H), 3.97–3.77 (m, 2H), 1.97 (s, 3H), 1.94 (s, 3H). ^{13}C NMR (100 MHz, *d$_6$*-DMSO): δ = 172.43, 172.18, 167.00, 157.31, 142.70, 140.25 (d, *J* = 5.4 Hz), 128.90 (d, *J* = 9.7 Hz) 121.66 (d, *J* = 5.5 Hz), 97.42, 88.73, 81.38 (d, *J* = 8.6 Hz), 74.07, 71.39, 65.84, 21.36, 21.24. ESI-HRMS: *m/z* [M + H]$^+$ calcd for C$_{16}$H$_{21}$N$_5$O$_9$P: 458.1071; found: 458.1070.

3. References

1   A. Maxut, B. Nozière, B. Fenet and H. Mechakra, *Phys. Chem. Chem. Phys.*, 2015, **17**, 20416–20424.

2   Y. Feng, D. W. Smith and J. R. Bolton, *J. Environ. Eng. Sci.*, 2007, **6**, 277–284.

3   Y.-Q. Wu, S. K. Hamilton, D. E. Wilkinson and G. S. Hamilton, *J. Org. Chem.*, 2002, **67**, 7553–7556.

4   P. B. W. McCallum, R. T. Weavers, M. R. Grimmett and A. G. Blackman, *Aust. J. Chem.*, 1999, **52**, 159–165.

5   L. Li, N. Prywes, C. P. Tam, D. K. O'Flaherty, V. S. Lelyveld, E. C. Izgu, A. Pal and J. W. Szostak, *J. Am. Chem. Soc.*, 2017, **139**, 1810–1813.

6   T. J. Henderson and D. B. Cullinan, *Magn. Reson. Chem.*, 2007, **45**, 954–961.

7   C. Na and T. M. Olson, *Environ. Sci. Technol.*, 2004, **38**, 6037–6043.

8   N. Borduas, B. Place, G. R. Wentworth, J. P. D. Abbatt, and J. G. Murphy, *Atmos. Chem. Phys.,* 2016*,* **16***,* 703–714.

9   C. Fernández-García and M. W. Powner, *Synlett*, 2017, **28**, 78–83.