Common and Potentially Prebiotic Origin for Precursors of Nucleotide Synthesis and Activation

Albert C. Fahrenbach, Constantin Giurgiu, Chun Pong Tam, Li Li, Yayoi Hongo, Masashi Aono, and Jack W. Szostak

SUPPORTING INFORMATION

Contents

1. General Methods

1a. General considerations. The preparatory synthesis of 2-aminoimidazole was conducted in air-dried round-bottomed flasks that were fitted with rubber septa. Reactions on 1 mL scales were all carried out in Eppendorf tubes. All reactions were carried out in 18 MΩ water using a MilliQ purification system. All reagents were purchased from Sigma Aldrich or TCI and used without further purification. Thin-layer chromatography (TLC) was performed using glass-backed precoated high-performance silica plates with blue fluorescent indicators (EMD Millipore, Billerica, MA), with detection by 254 nm ultraviolet light and/or by alkaline potassium permanganate stains. Normal-phase flash chromatography was carried out with technical grade silica gel (60 Å pore size, 230–400 mesh particle size). D_2O for NMR analysis was purchased from Cambridge Isotope Laboratories (Tewksbury, MA). Aqueous solutions of cytidine 5´-monophosphate (disodium salt form, CMP•2Na) in D_2O were used as internal standards for integration, and their concentrations were determined by their UV absorption at 271 nm on a Thermo Scientific Nanodrop 2000c Spectrophotometer (Waltham, MA), using a molar extinction coefficient of 9000 M^{-1} cm^{-1}.

1b. NMR spectroscopy. NMR spectra were acquired on either a Varian INOVA 400 MHz or Bruker Avance^{III} 400 MHz NMR spectrometer equipped with a broadband PFG (z-gradient) probe and a 5 mm z-gradient PABBO BB-1H/D Z-GRD probe, respectively (400 MHz for ^{1}H, 100 MHz for ^{13}C, 161 MHz for ^{31}P). Proton, phosphorus and carbon chemical shifts are reported in parts per million (ppm) values on the δ scale, and internally referenced to residual signals in the NMR solvent unless otherwise noted (^{1}H: DHO, $\delta = 4.79$ ppm, $CHCl_3$ $\delta = 7.26$ ppm; ^{13}C: $CDCl_3 = 77.16$ ppm). All NMR spectra were recorded at 25 ^{o}C and were analyzed using MestReNova (MestreLab

Research, Santiago de Compostela, Spain). Baseline subtraction was carried out before all resonance integrations.

1c. Low-resolution mass spectroscopy (LRMS). Low-resolution mass spectrometry (LRMS) analyses were performed using a Bruker Daltonics Esquire 6000 mass spectrometer.

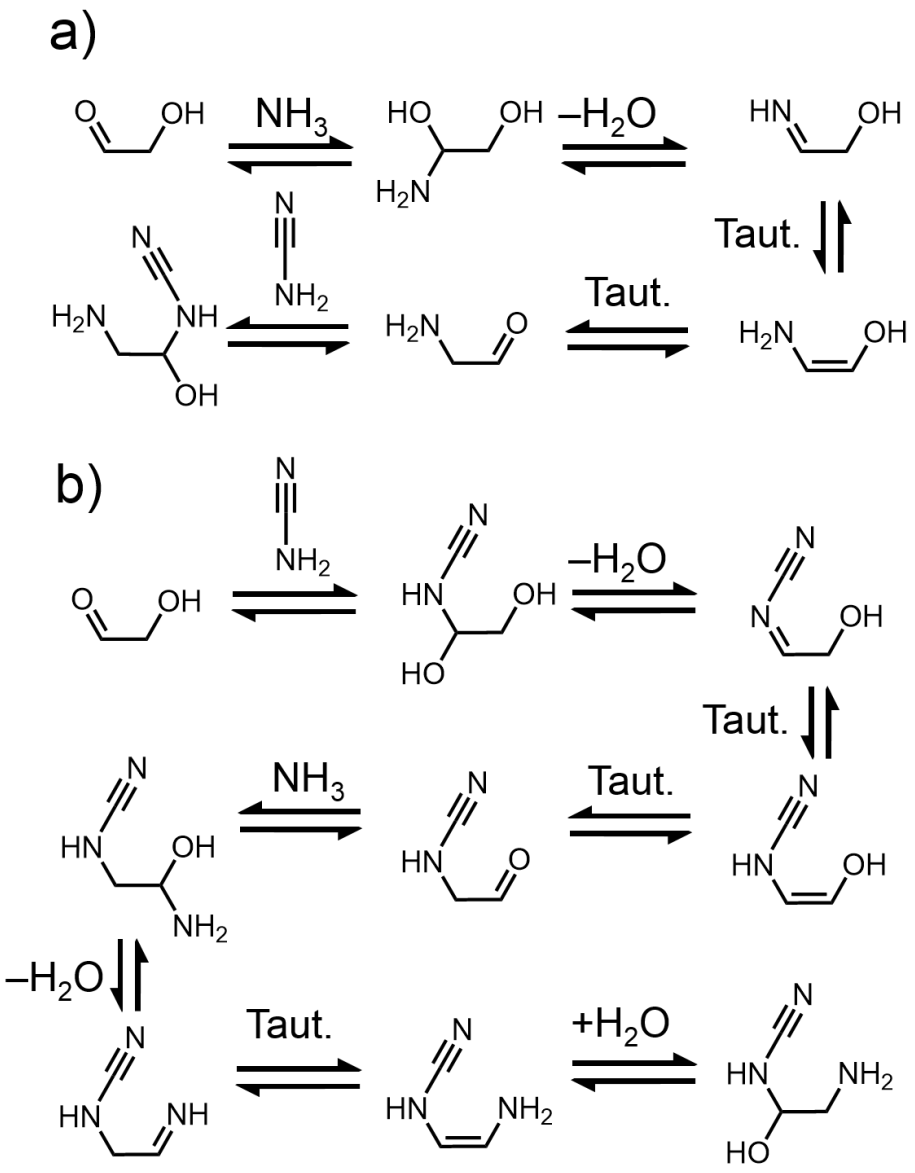
1d. High-resolution mass spectrometry (HRMS). High-resolution mass data analyses were performed on a Xevo-Q-TOF-MS platform (Waters, Manchester, UK) combined with a UPLC system (Acquity; Waters, MA). Data analysis was performed using MestReNova. Liquid chromatography was carried out on a CORTECS UPLC C_{18} column (1.6 *u*m, 2.1 mm x 150 mm, Waters) at a flow rate of either 0.2 or 0.3 mL min^{-1}. Solvent A was 0.1% formic acid in water and solvent B was acetonitrile. **2NH$_2$Im**, **2NH$_2$Ox** and the mixture of aminooxazolines all eluted in 100% A before two minutes when the flow rate was 0.3 mL min^{-1}. Solvent B was then ramped from 0 to 90% from 2.0 to 2.1 minutes and remained at 90% until 4 minutes. Solvent B was then ramped back down to 0% at 4.1 minutes, and 100% solvent A was eluted for an additional two minutes. The column was heated to 30 or 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 0.5 *u*L. Under positive mode ESI conditions, a voltage of 3.0 kV was applied to the stainless-steel electrospray ionization. The TOF analyzer was set to sensitivity mode with a resolving power of 22,000, and the *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.

2. Supplementary Discussion

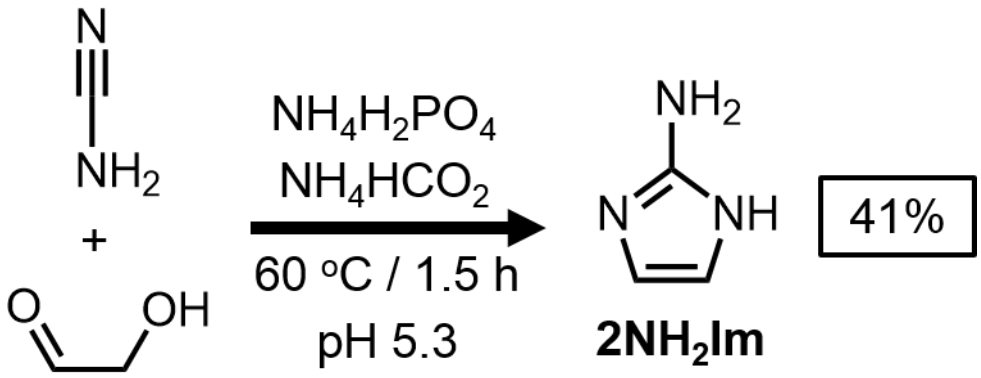
Mechanisms of Prebiotic Imidazole Synthesis. Several plausible mechanisms for the prebiotic synthesis of imidazoles have been reported[1]. Oró and coworkers[2] reported a pathway that employs the condensation of glyoxal with formaldehyde or acetaldehyde in the presence of ammonia. This mechanism allows for the synthesis of imidazole and 2-methylimidazole, respectively. Another mechanism also demonstrated by Oró et al.[3] employs condensation of erythrose with formamidine and ammonia to yield imidazole-4-acetaldehyde and imidazole-4-glycol. The even simpler reaction of racemic glyceraldehyde with aqueous ammonia reported by Grimmett and Richards[4] leads to a distribution of 4-substituted imidazoles. A very different mechanism for the prebiotic synthesis of imidazoles that avoids aldehyde precursors involves the UV-light promoted photocyclization of enaminonitriles[5,6]. A notable example is the cyanide tetramer, diaminomalononitrile, which Ferris and Orgel[7] showed undergoes conversion upon photoexcitation with 253 nm light to 4-aminoimidazole-5-carbonitrile, an intermediate *en route* to adenine and other purines.

Mechanisms of Prebiotic Ammonia Production. There are several scenarios in which ammonia could have been generated in a concentrated form within an ancient aqueous reservoir. One source of ammonia could be the reduction of aqueous cyanide by solvated electrons[8], generating methanimine which in turn hydrolyzes to formaldehyde and ammonia. Alternatively, ammonia gas formed in the atmosphere could partition into the aqueous reservoir, and at mildly acidic pH protonation to ammonium would push the equilibrium away from that predicted by Henry's Law and towards the aqueous phase. Other pathways could involve the UV-light activated photoreduction of N_2 by the mineral rutile[9] or reduction of nitrite[10] by aqueous Fe(II).

3. Supplementary Schemes



Scheme S1. Possible Amadori-type mechanisms for exchange of water for ammonia a) before and b) after addition of cyanamide to the carbonyl of glycolaldehyde.



Scheme S2. Prep-Scale Synthesis of 2-Aminoimidazole (**$2NH_2Im$**).

2-Aminoimidazole ($2NH_2Im$). A solution containing $NH_4H_2PO_4$ (0.96 g, 8.3 mmol) and NH_4HCO_2 (2.60 g, 41 mmol) in 5 mL of H_2O was prepared. The resulting pH of this solution is ~ 5.3. To this solution was added glycolaldehyde (0.50 g, 8.3 mmol) and cyanamide (0.35 g, 8.3 mmol) directly as solids. After addition of the solids, the total volume was approximately 8.3 mL. The mixture was allowed to react for 1.5 hours at 60 °C. The reaction mixture was removed from heat and allowed to cool to room temperature before adding 3.0 mL of 50% NaOH followed by addition of 100 mL of MeOH while stirring rapidly. The resulting white precipitate of inorganic phosphate was filtered off immediately, washed with additional MeOH and discarded. The filtrate and washings were combined without hesitation and concentrated to dryness under reduced pressure. The resulting crude, orange oil was purified by column chromatrography (SiO_2, solid loading; CH_2Cl_2 / 1 N NH_3 in MeOH). Eluting with 30% 1 N NH_3 in MeOH removed a broad, light orange band, which was the target product in pure from. After the solvent was evaporated *in vacuo*, 0.28 g of a light orange oil was collected, in 41% yield (LRMS *m/z* = 84.3; calculated: 84.1 for $C_3N_3H_6$).

4. Supplementary Figures

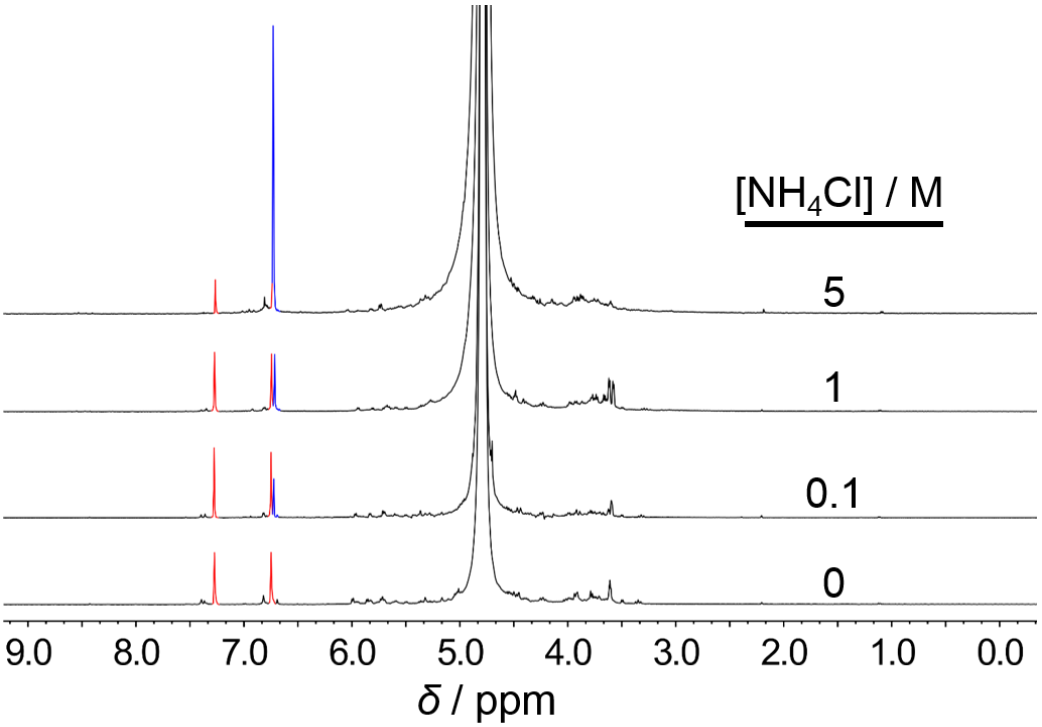


Figure S1. The yield of **2NH$_2$Im** increases with higher concentrations of NH_4Cl. Glycolaldehyde (1 M) and cyanamide (1 M) were reacted with a series of concentrations of ammonium chloride, in the presence of sodium phosphate (1 M). The spectra shown were all recorded after reacting for 3 hours at 60 ^oC and pH 7. Each reaction was carried out on a 1 mL scale. Reactions were monitored over time. For each time point, a 100 uL aliquot was sampled from the reaction mixture, diluted into 350 *u*L of D_2O, and the ^1H NMR spectrum recorded. Concentrations of **2NH$_2$Im** and **2NH$_2$Ox** were measured by adding 50 *u*L of a calibrated solution of CMP•2Na to the NMR tube. Resonances for **2NH$_2$Ox** are colored in red; the resonance for **2NH$_2$Im** is colored in blue. Resonance assignments were confirmed by additions of authentic standards.

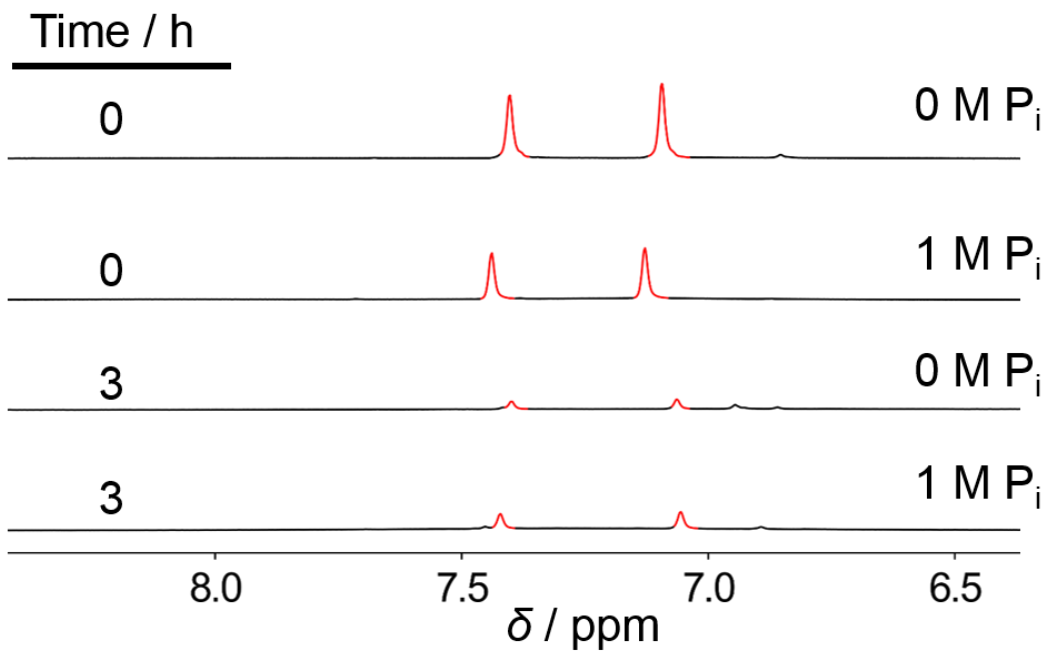


Figure S2. **2NH$_2$Ox** does not convert to **2NH$_2$Im** when treated with NH_4Cl in the presence of phosphate. The spectra shown were recorded before and after reacting for 3 hours at 60 ^oC and pH 4, with all reagents at 1 M. Each reaction was carried out on a 1 mL scale. Reactions were monitored over time. For each time point shown, a 100 uL aliquot was sampled from the reaction mixture and was diluted into 350 *u*L of D_2O, and the ^1H NMR spectrum was recorded. Concentrations of **2NH$_2$Ox** were measured by adding 50 *u*L of a calibrated solution of CMP•2Na to the NMR tube. Resonances for **2NH$_2$Ox** are colored in red. Resonance assignments were confirmed by additions of authentic standards. No signals arising from **2NH$_2$Im** could be observed after 3 hours, while the signals for **2NH$_2$Ox** decreased in intensity. We hypothesize that this decrease in intensity is a result of the fact that some of the **2NH$_2$Ox** is in equilibrium with its hydrate at this lower pH.

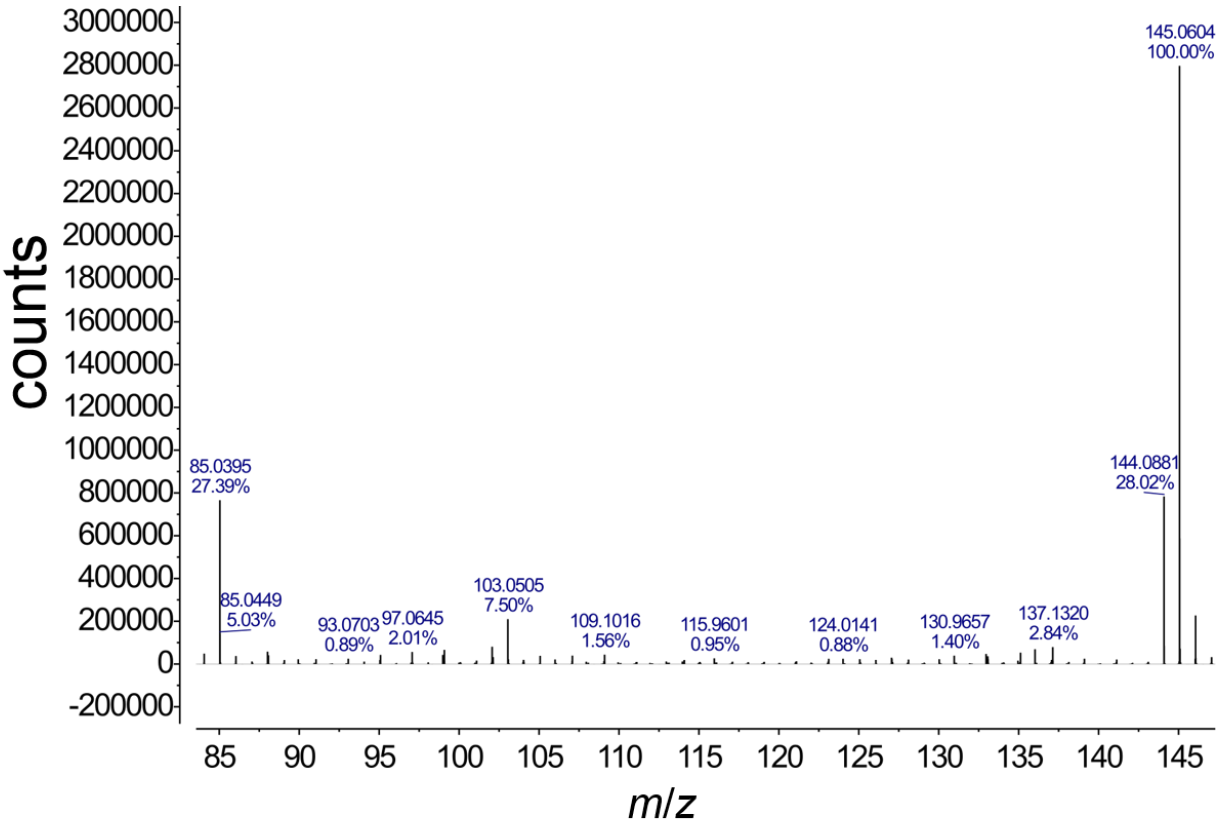


Figure S3. Partial mass spectrum recorded from a sample of the reaction described in Table S2 that employed 0 M NH_4Cl at pH 7 after 1.75 h. The *m*/z value of 145.0604 is consistent with an addition product of **2NH$_2$Ox** and glycolaldehyde, a side-product which may help account for the relatively low yields observed. This peak is larger in intensity compared to the *m*/*z* = 85.0395 peak corresponding to **2NH2Ox**, an observation which may be related to the relatively high nucleophilicity of **2NH$_2$Ox.**

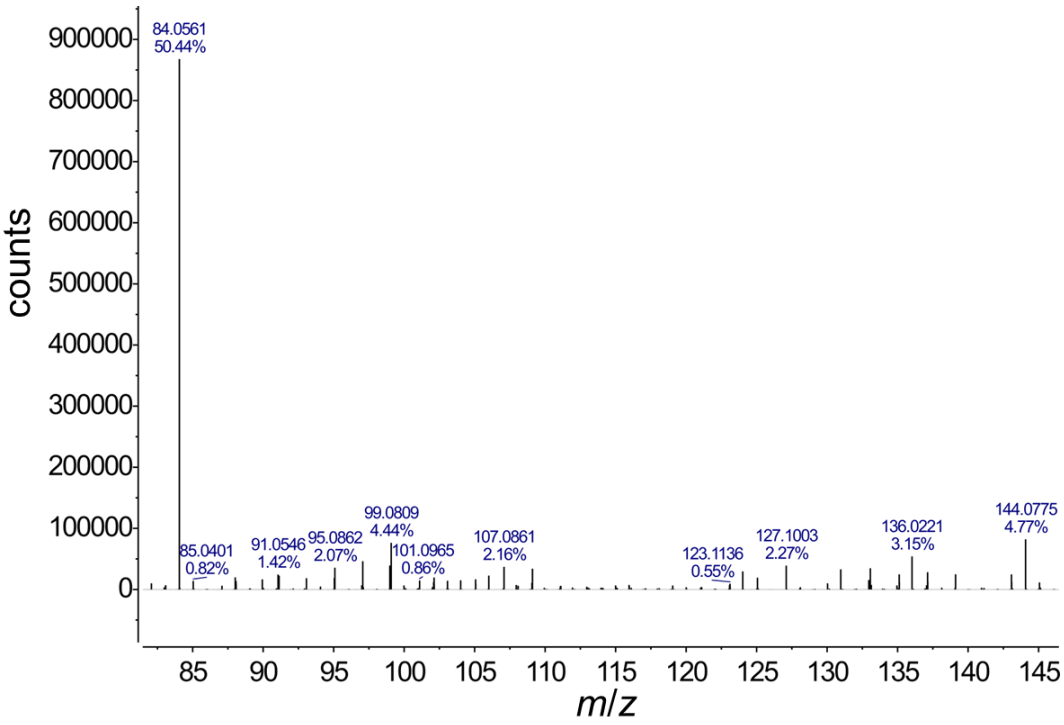


Figure S4. Partial mass spectrum recorded from the reaction described in Table S2 at 5 M NH_4Cl and pH 5.5 at 1.75 h. The *m*/*z* value of 144.0775 peak is consistent with an addition product of **2NH$_2$Im** and glycolaldehyde, another side-product which may help account for the relatively low yields observed. However, the intensity of this peak is much less in comparison to the *m*/*z* = 84.0561 peak corresponding to **2NH$_2$Im**, an observation which hints at the relatively lesser nucleophilicity of the **2NH$_2$Im** ring.

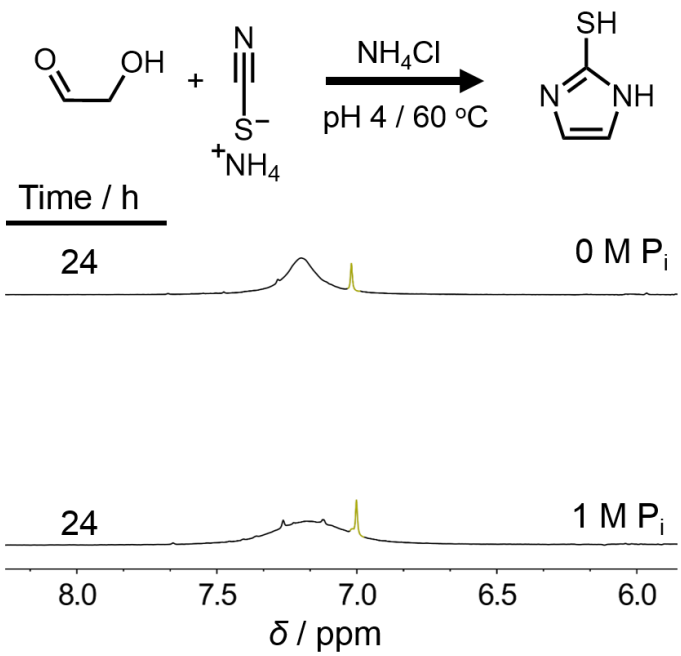


Figure S5. Prebiotic synthesis of 2-thioimidazole. Glycolaldehyde (1 M) and ammonium thiocyanate (1 M), were reacted with ammonium chloride (4 M) with and without sodium phosphate (1 M). The spectra shown were recorded before and after reacting for 24 hours at 60 ^{o}C and pH 4. Each reaction was carried out on a 1 mL scale. The reactions were monitored over time. For the time points shown, a 100 uL aliquot was sampled from the reaction mixture and was diluted into 350 *u*L of D_2O, and the ^{1}H NMR spectrum was recorded. Concentrations of 2-thioimidazole were measured by adding 50 *u*L of a calibrated solution of CMP•2Na to the NMR tube. The resonance arising from the H4 and H5 protons of 2-thioimidazole is colored in gold and its assignment was confirmed by additions of an authentic standard. The broad peak adjacent to the 2-thioimidazole resonance is believed to arise from ammonium protons in intermediate exchange with those of H_2O. This resonance no longer appears after addition of CMP•2Na. The presence of 1 M phosphate slightly increases the yield of 2-thioimidazole, from 6.2% to 7.8%.

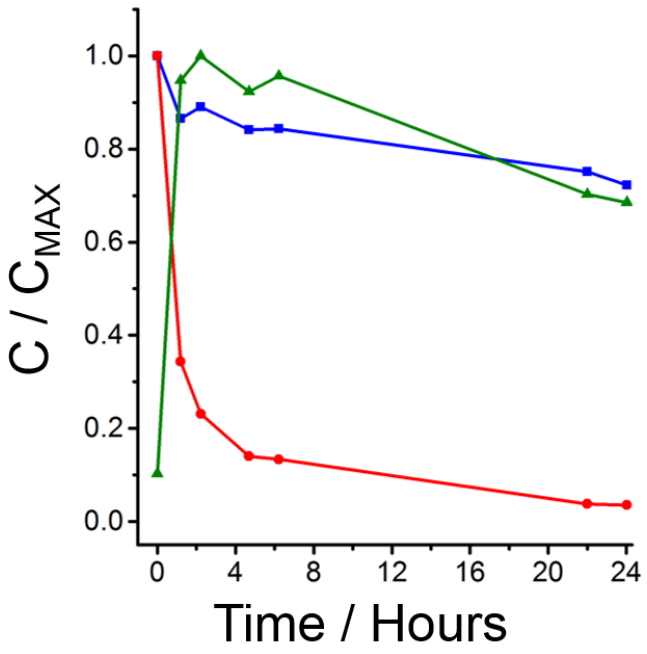


Figure S6. **2NH$_2$Im** does not interfere with the reaction of *rac*-glyceraldehyde and **2NH$_2$Ox**. The reaction was carried out at a 1 mL scale in the presence of sodium phosphate (1 M) at pH 7 and 40 ^{o}C over a period of 24 hours. The reaction was monitored over time using Q-TOF LCMS on a C_{18} column. For the time points shown, a 10 uL aliquot was sampled from the reaction mixture and then diluted into 1 mL of H_2O and frozen at –20 ^{o}C. After all aliquots had been collected and frozen, they were then thawed and samples for LCMS were prepared by a further 100-fold dilution. In the graph above, C/C_{MAX} represents the normalized relative abundances for **2NH$_2$Ox** (red), **2NH$_2$Im** (blue) and the mixture of pentose 2-aminooxazolines (green) to each compounds observed maximum. C/C_{MAX} values were determined from integrations obtained from extracted ion chromatograms: **2NH$_2$Ox** (m/z = 85.04 ±0.1, calculated: 85.0402 for $[M + H]^+$), **2NH$_2$Im** (m/z = 84.06 ±0.1, calculated: 84.0562 for $[M + H]^+$), pentose 2-aminooxazoline (m/z = 175.07 ± 0.1, calculated: 175.0719 for $[M + H]^+$).

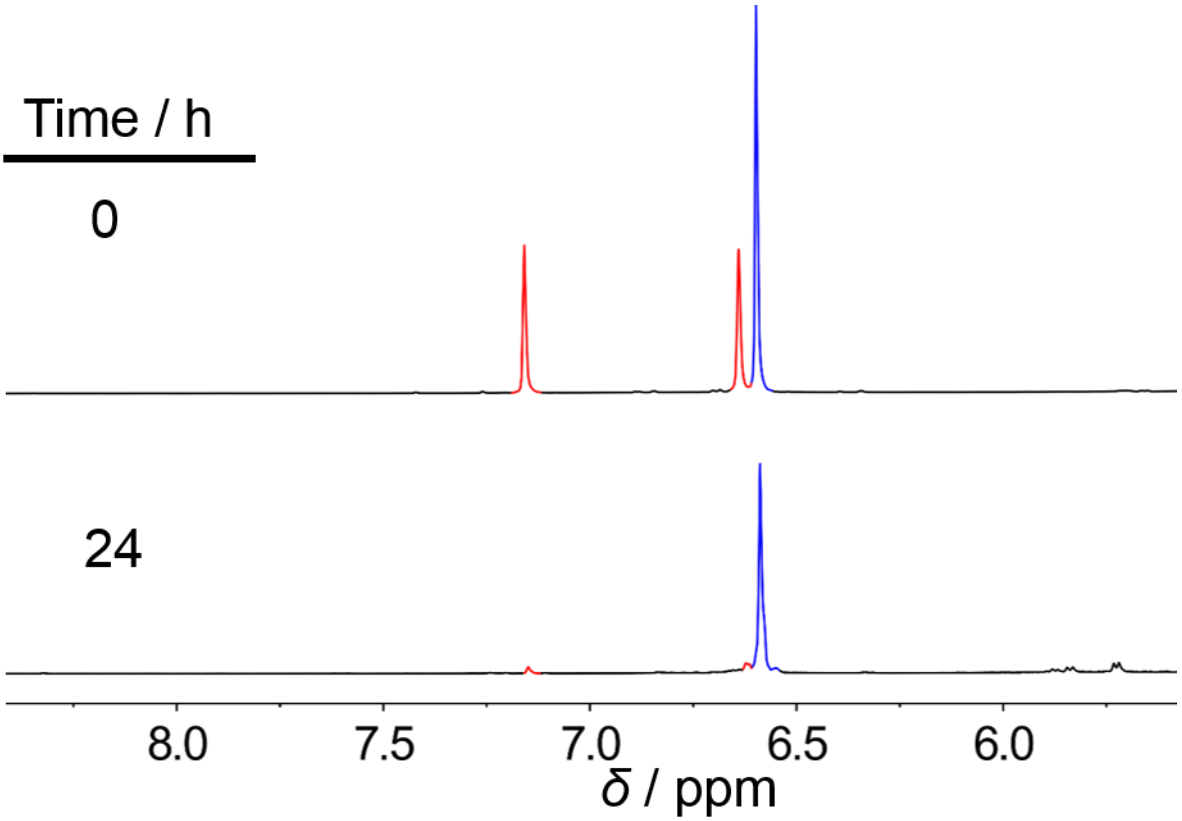


Figure S7. Glyceraldehyde selectively reacts with **2NH$_2$Ox** in the presence of **2NH$_2$Im**. ^{1}H NMR spectra of reactions of *rac*-glyceraldehyde (1 M), **2NH$_2$Ox** (1 M), and **2NH$_2$Im** (1 M) with sodium phosphate as a catalyst (1 M). The spectra shown were recorded before and after reacting for 24 hours at 40 ^{o}C and pH 7. The reaction was carried out on a 1 mL scale. For the time points shown, a 100 uL aliquot was sampled from the reaction mixture and was diluted into 350 *u*L of D_2O, and the ^{1}H NMR spectrum was recorded. Resonances for **2NH$_2$Ox** are colored in red, while those for **2NH$_2$Im** are in blue. After 24 hours, the resonances for **2NH$_2$Ox** have nearly gone to zero in intensity, consistent with the results obtained from LCMS wherein **2NH$_2$Ox** reacts with *rac*-glyceraldehyde and then cyclizes to form a mixture of pentose aminooxazolines.

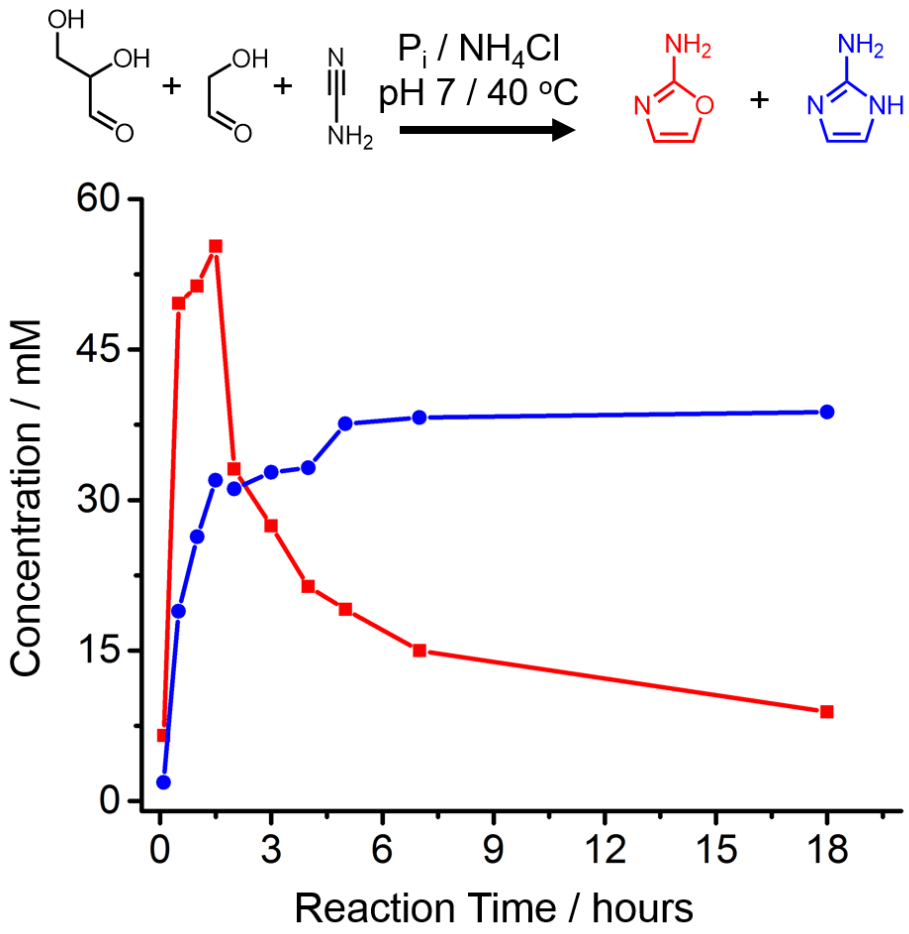


Figure S8. **2NH$_2$Im** is synthesized in a one-pot reaction of 1 M cyanamide, glycolaldehyde, glyceraldehyde in 1 M sodium phosphate and 1 M NH_4Cl at pH 7 and 40 °C. The reaction was monitored over time by Q-TOF LCMS on a C_{18} column. Concentrations of **2NH$_2$Ox** and **2NH$_2$Im** were determined from a calibration curve made from authentic standards. The concentration of **2NH$_2$Ox** rapidly increases, reaches a maximum after 1.5 hours, and subsequently begins to decline, presumably at least in part due to the formation of pentose aminooxzolines. The **2NH$_2$Im** concentration undergoes an initial rapid increase and levels off after about 5 hours. This experiment demonstrates that even in the presence of glyceraldehyde, **2NH$_2$Im** can still form and accumulate to a significant amount.

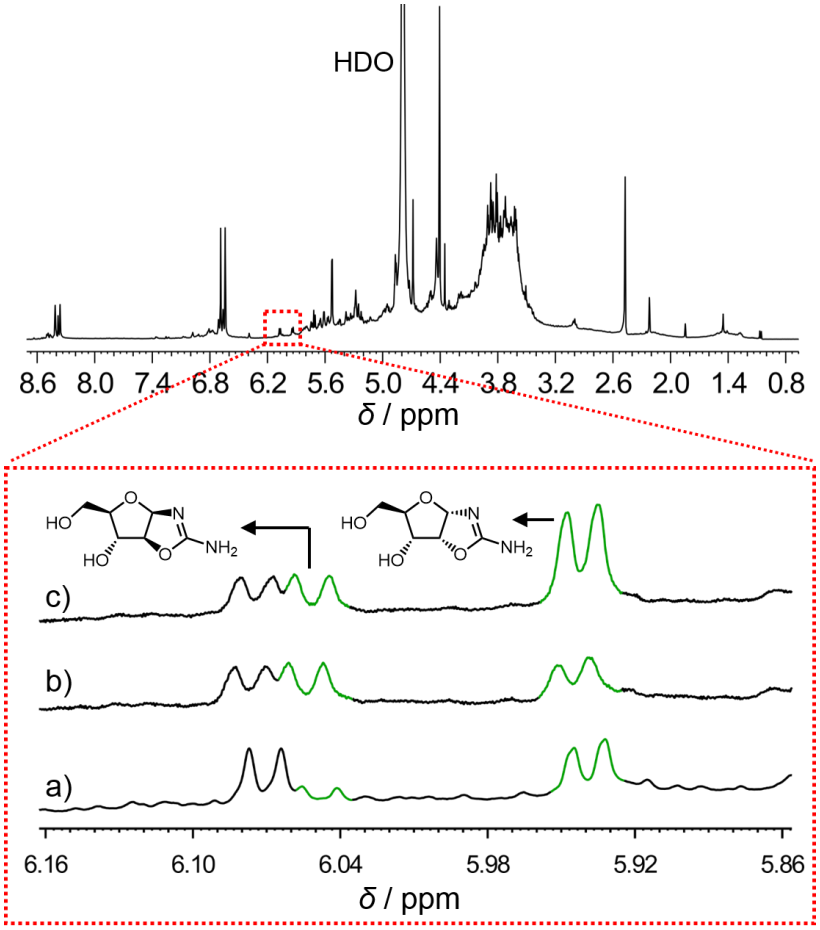


Figure S9. Pentose aminooxazolines are synthesized in the one-pot reaction of 1 M cyanamide, glycolaldehyde, glyceraldehyde with 1 M sodium phosphate and 1 M NH_4Cl at pH 7 and 40 °C. A sample of the same reaction mixture as the experiment described in Figure S8 was analyzed by ^1H NMR spectroscopy after 18 hours (top spectrum). The spectrum was recorded after a 300 uL aliquot of the reaction mixture was first dried under reduced pressure, and then the residual solid was re-dissolved in D_2O. Authentic standards of D-ribose aminooxazoline and D-arabinose aminooxazoline were prepared[11] according to previously reported procedures. a) Partial ^1H NMR spectrum zoomed-in on the 6 ppm region; b) after addition of D-arabinose aminooxazoline; c) after addition of D-ribose aminooxazoline. The proton resonances colored in green are observed to increase in intensity after addition of the standards. These resonances arise from the C1 protons of each compound, respectively.

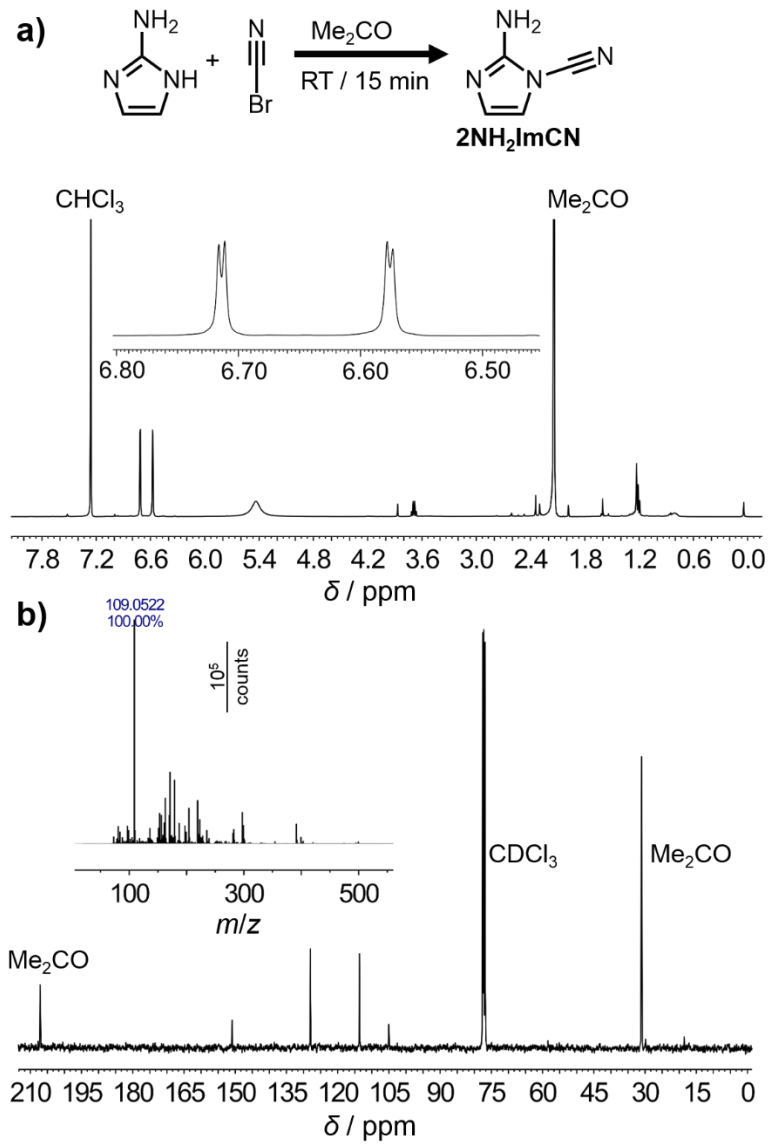


Figure S10. Synthetic scheme for **2NH$_2$ImCN**. **2NH$_2$Im** (50 mg, 0.6 mmole) was dissolved in 6 mL of acetone. A solution of BrCN (70 mg, 0.66 mmole) in 1 mL of acetone was added. The reaction was allowed to stir at room temperature for 15 minutes before it was diluted with 100 mL of $CHCl_3$. The mixture was filtered, and the filtrate was concentrated to near dryness under vacuum at 35 ^oC. When only a small volume of liquid remained, the vacuum was removed and 1 mL of $CDCl_3$ was added, and 0.5 mL of this mixture was transferred an NMR tube for analysis. a) ^1H NMR spectrum of the transferred solution. The inset shows the partial ^1H NMR spectrum from ~6.8 to 6.5 ppm, and displays the two resonances arising from the H4 and H5 protons of **2NH$_2$ImCN**. b) ^{13}C NMR spectrum of the same solution. The inset shows the mass spectrum taken of a reaction performed in a similar manner and reveals a peak with an *m*/*z* value of 109.0522, which is consistent with the target product (calc. for $[M + H]^+$: 109.051). Once the compound is concentrated to dryness, decomposition begins to set in almost immediately. Direct ^1H NMR spectroscopic analysis of the reaction mixture after 15

minutes was carried out by taking a 250 *u*L aliquot of the reaction mixture and combining with 250 *u*L d_6-DMSO. The percent conversion based on relative integration of the signals arising from **2NH$_2$ImCN** and **2NH$_2$Im**, respectively, was measured to be 43%. No significant change in the conversion was observed after 1 hour of reaction time.

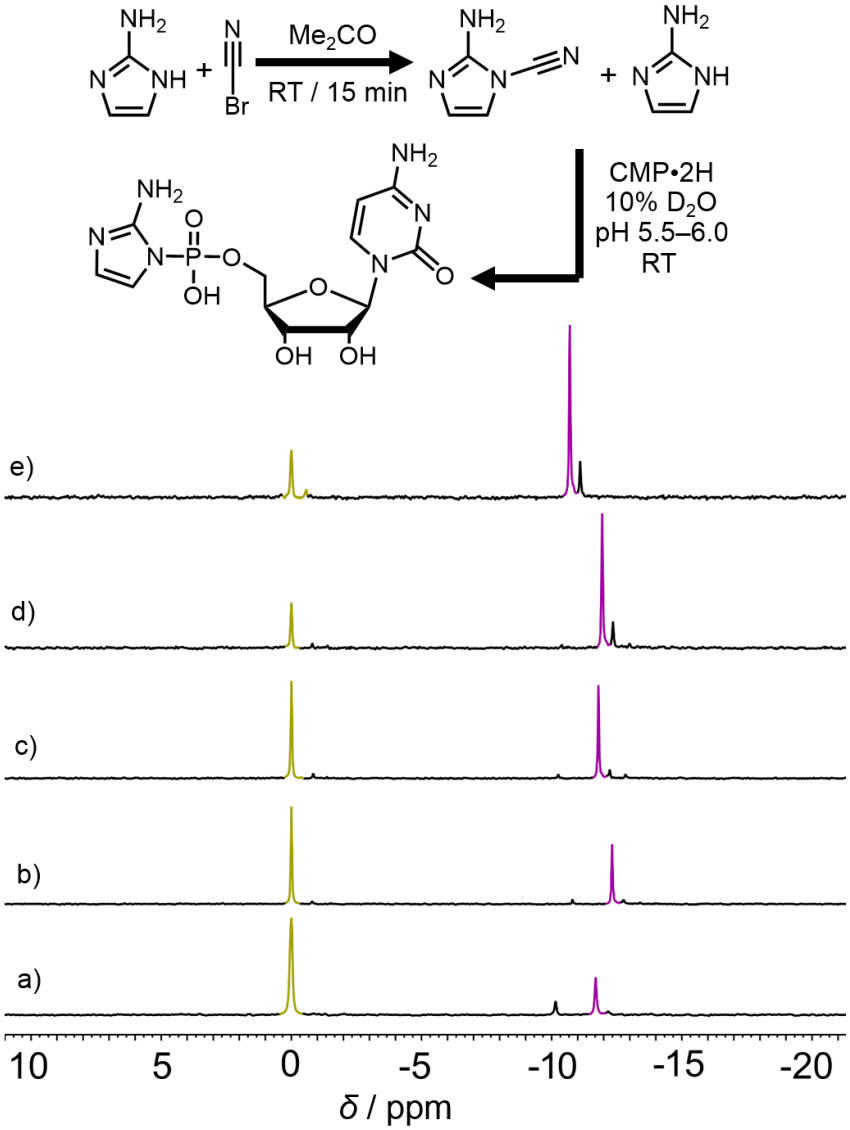


Figure S11. Synthesis of cytidine-5′-phosphoro-(2-aminoimidazole) (**2NH$_2$ImpC**). The reaction was carried out as follows. **2NH$_2$Im** (21 mg, 0.25 mmole) was dissolved in 3 mL of acetone. A solution of cyanogen bromide (16 mg, 0.15 mmole) in 1 mL of acetone was added. The reaction was allowed to stir for 15 minutes at room temperature before being concentrated to a minimal volume under vacuum at 35 ^{o}C. As a consequence of the fact that **2NH$_2$ImCN** immediately begins to undergo decomposition when isolated as a solid, the crude reaction mixture was used as is without further purification. An aqueous solution of 5′-cytidylic acid (5 mg, 0.015 mmole) in 0.45

mL of H_2O was added, and the mixture was kept under vacuum for a few additional minutes to remove the remainder of the acetone. The initial pH of the reaction was 5.5. The reaction was transferred to an NMR tube along with 50 uL of D_2O. The pH of the reaction was monitored, and it tended to increase over time. Concentrated HCl was used to keep the pH of the reaction between 5.5 and 6.0. Parial ^{31}P NMR spectrum of the reaction mixture after a) 10 minutes b) 40 minutes and c) 1 h and 40 minutes. A freshly prepared solution of **2NH$_2$ImCN** was then prepared in the same way as described above and combined with contents of the NMR tube, and the ^{31}P NMR spectrum was recorded d) after an additional 30 minutes and e) after 1 hour and 15 minutes. The resonance colored in yellow is assigned to the phosphorus atom of the starting material, which was referenced to 0 ppm, while the resonance colored in purple is assigned to the phosphorus of the target **2NH$_2$ImpC**. The changes in chemical shift during the reaction are thought to be a consequence of the changing pH of the solution.

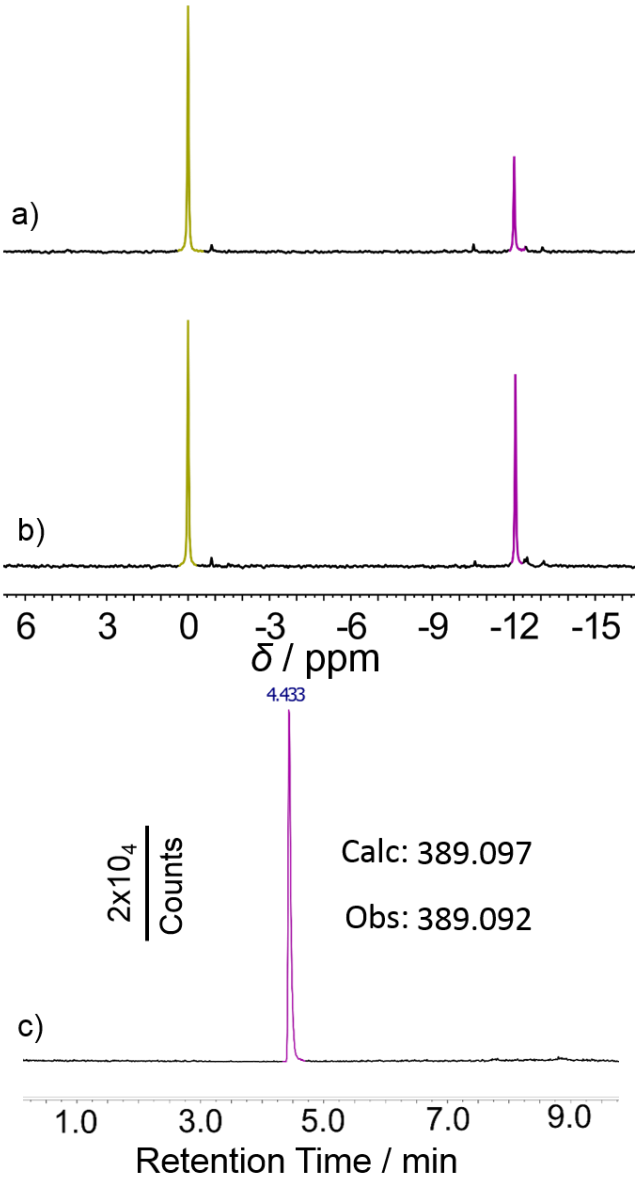


Figure S12. Confirmation of the synthesis of **2NH$_2$ImpC**. The same synthetic procedure described above in Figure S11 was carried out. a) ^{31}P NMR spectrum of the reaction mixture after 45 minutes and b) after addition of authentic standard of **2NH$_2$ImpC** synthesized by a previously reported procedure[12]. After addition of the authentic standard, the resonance colored in purple increases in intensity, confirming that it arises from the phosphorus atom of **2NH$_2$ImpC**. c) Extracted ion chromatogram for the $[M + H]^+$ mass of **2NH$_2$ImpC**. Q-TOF LCMS was carried out on the same reaction mixture as in a).

pH→ Time↓ h	4			5.5			7			8.5		
	$2NH_2Ox$ % Yield	**$2NH_2Im$** % Yield	**$[2NH_2Im]/[2NH_2Ox]$**	**$2NH_2Ox$** % Yield	**$2NH_2Im$** % Yield	**$[2NH_2Im]/[2NH_2Ox]$**	**$2NH_2Ox$** % Yield	**$2NH_2Im$** % Yield	**$[2NH_2Im]/[2NH_2Ox]$**	**$2NH_2Ox$** % Yield	**$2NH_2Im$** % Yield	**$[2NH_2Im]/[2NH_2Ox]$**
0 M [NH_4Cl]												
1	4.2	0.8	0.18	10.1	0.9	0.09	12.5	0.3	0.021	9.5	0.1	0.014
2	2.6	0.9	0.35	6.7	0.8	0.12	14.8	0.8	0.053	9.9	0.3	0.026
3	2.6	1.0	0.40	4.9	0.5	0.11	10.7	0.5	0.049	10.1	0.4	0.038
0.5 M [NH_4Cl]												
1	1.3	5.9	4.5	3.1	5.3	1.7	13.3	3.0	0.23	11.4	1.7	0.15
2	1.3	6.4	4.9	4.2	5.7	1.4	11.4	2.9	0.25	10.7	1.8	0.17
3	1.3	6.2	4.8	3.1	5.5	1.8	9.8	3.0	0.31	10.1	2.0	0.19
1 M [NH_4Cl]												
1	2.1	9.6	4.6	1.8	9.8	5.4	9.6	6.6	0.69	22.4	11.6	0.52
2	1.0	10.3	9.9	1.6	10.5	6.8	9.4	7.0	0.75	8.6	5.2	0.61
3	1.3	10.3	7.9	1.3	10.0	7.7	8.8	7.3	0.82	5.5	4.8	0.88
2 M [NH_4Cl]												
1	1.2	13.9	12.0	1.8	15.9	8.8	10.8	14.4	1.3	12.2	12.1	1.0
2	1.3	14.6	11.2	3.0	14.8	5.0	6.6	16.6	2.5	8.5	13.1	1.5
3	1.4	14.4	10.7	0.7	15.9	23.6	6.9	16.6	2.4	7.3	13.3	1.8
3 M [NH_4Cl]												
1	0.6	12.8	21.3	1.0	15.9	16.6	5.3	14.8	2.8	6.0	10.1	1.7
2	0.8	14.1	17.1	0.6	14.7	26.3	4.0	14.2	3.6	4.3	9.9	2.3
3	1.7	22.2	13.3	0.4	16.0	40.0	2.0	14.0	7.0	2.7	8.1	3.0
4 M [NH_4Cl]												
1	0.5	12.3	27.3	0.7	16.7	23.0	2.7	16.5	6.1	4.0	11.4	2.9
2	0.4	14.9	37.3	0.3	16.6	51.1	1.3	15.9	12.2	2.7	10.1	3.8
3	0.4	15	37.5	0.3	15.5	51.7	1.1	16.5	15	1.1	7.9	7.2
5 M [NH_4Cl]												
1	0.3	9.9	38	0.5	13.4	26	1.0	15.2	15	5.2	8.7	1.7
2	0.5	11.3	22	0.3	13.8	53	0.8	14.8	19	2.3	9.8	4.2
3	0.5	11.4	22	0.3	14.4	56	0.5	14.6	28	1.0	8.5	8.1

Table S1. Yields and ratios of **2NH$_2$Im** to **2NH$_2$Ox** at variable pH and concentrations of ammonium chloride. Values shown here were used to generate the bar graph shown in Figure 2 of the main text. All reactions were carried out using glycolaldehyde (1 M), cyanamide (1 M) and sodium phosphate (1 M). First, the pH of the solutions containing only ammonium chloride and sodium phosphate were adjusted using concentrated solutions of NaOH or HCl, and then the reactions were initiated by adding cyanamide and glycolaldehyde to the solutions as solids. Each reaction condition was carried out on a 1 mL scale, and samples for analysis were taken at 1, 2 and 3 hour time points. For each time point, a 100 *u*L aliquot was sampled from the reaction mixture, was frozen at –80 ^{o}C, and then lyophilized at RT. After all the water had been removed, the resulting solids were dissolved in 450 *u*L of D_2O. Note, addition of NaOH and MeOH was only carried out in the case of the prep-scale reaction procedure. Concentrations of **2NH$_2$Im** and **2NH$_2$Ox** were measured by adding 50 *u*L of a calibrated solution of CMP•2Na in D_2O to the sample, and the solutions were transferred to NMR tubes and the ^{1}H NMR spectra were recorded. For most cases, the ratios of **2NH$_2$Im**/**2NH$_2$Ox** tended to increase over time. The reason for this increase is not the increase in concentration of **2NH$_2$Im** after one hour, but rather the decrease in the yield of **2NH$_2$Ox**. This decrease in the yield observed over time for most conditions tested is likely a reflection of the greater reactivity of **2NH$_2$Ox**, possibly forming a distribution of side-products most of which are not easily identified by ^{1}H NMR.

pH→		4			5.5			7			8.5	
Time h↓	**$2NH_2Ox$** %Yield	**$2NH_2Im$** %Yield	**[$2NH_2Im$] / [$2NH_2Ox$]**	**$2NH_2Ox$** %Yield	**$2NH_2Im$** %Yield	**[$2NH_2Im$] / [$2NH_2Ox$]**	**$2NH_2Ox$** %Yield	**$2NH_2Im$** %Yield	**[$2NH_2Im$] / [$2NH_2Ox$]**	**$2NH_2Ox$** %Yield	**$2NH_2Im$** %Yield	**[$2NH_2Im$] / [$2NH_2Ox$]**
0 M [NH_4Cl]												
1.75	19.7	2.9	0.1	29.9	2.6	0.1	28.8	1.5	0.1	23.0	0.7	0.0
3.5	3.8	1.8	0.5	6.9	1.4	0.2	11.4	1.0	0.1	15.8	0.6	0.0
0.5 M [NH_4Cl]												
1.75	11.4	15.3	1.3	19.1	13.6	0.7	19.4	5.0	0.3	29.2	6.1	0.2
3.5	2.9	10.8	3.7	4.0	9.5	2.4	10.8	6.0	0.6	12.3	4.8	0.4
1 M [NH_4Cl]												
1.75	6.7	16.3	2.4	11.1	16.8	1.5	22.4	11.1	0.5	26.6	10.6	0.4
3.5	1.5	14.9	9.8	1.9	13.5	7.2	10.2	11.2	1.1	11.5	7.5	0.6
2 M [NH_4Cl]												
1.75	2.0	20.3	10.0	4.7	20.8	4.5	9.9	17.2	1.7	12.2	12.7	1.0
3.5	1.2	16.0	13.5	1.8	16.8	9.3	5.9	14.0	2.4	14.4	6.1	0.4
3 M [NH_4Cl]												
1.75	1.4	16.1	11.6	3.4	19.3	5.8	7.3	16.9	2.3	9.4	13.6	1.4
3.5	0.62	15.7	25.2	1.6	18.1	11.0	3.3	17.6	5.4	3.3	17.6	5.4
4 M [NH_4Cl]												
1.75	0.96	16.8	17.6	1.8	18.0	9.8	3.9	19.2	4.9	6.3	12.8	2.0
3.5	0.43	19.7	45.6	0.58	20.2	34.9	2.2	20.6	9.4	2.7	12.0	4.4
5 M [NH_4Cl]												
1.75	1.5	18.8	12.3	2.8	27.3	9.8	5.5	26.7	4.9	8.0	20.6	2.6
3.5	0.45	17.1	38.4	0.5	21.8	43.3	1.0	20.5	19.7	2.7	13.4	5.0

Table S2. Yields and ratios of **2NH$_2$Im** to **2NH$_2$Ox** at different pH values and concentrations of ammonium chloride measured by Q-TOF LCMS (positive mode). All reactions were carried out using glycolaldehyde (1 M), cyanamide (1 M) and sodium phosphate (1 M). First, the pH of the solutions containing only ammonium chloride and/or sodium phosphate were adjusted using concentrated solutions of NaOH or HCl, and then cyanamide and glycolaldehyde were added to the solutions as solids. Each reaction condition was carried out on a 1 mL scale, and samples for analysis were taken at 1.75 and 3.5 hour time points. For each time point, a 10 *u*L aliquot was sampled from the reaction mixture, diluted with 990 *u*L of H_2O and then frozen at –20 ^oC. After all time points were collected, the samples were thawed, diluted an additional 100-fold, and transferred to vials for mass spectrometry and placed in an autosampler cooled to 4 ^oC. Note, addition of NaOH and MeOH was only carried out in the case of the prep-scale reaction procedure. Concentrations of **2NH$_2$Im** and **2NH$_2$Ox** were measured from a calibration curve determined from authentic standards. The yields and ratios of [**2NH$_2$Im**]/[**2NH$_2$Ox**] determined by LCMS are consistent with those determined by NMR (Table S1). Other than the *m*/*z* values arising from **2NH$_2$Im** and **2NH$_2$Ox**, we also observed a major peak at *m*/*z* = 145.06 (Fig. S3), and a minor peak at *m*/*z* = 144.08 (Fig. S4), consistent with the addition products of glycolaldehyde with **2NH$_2$Ox** and **2NH$_2$Im**, respectively. These potential addition products could be either the hemiaminal obtained from the attack of the exocyclic amines of **2NH$_2$Ox** or **2NH$_2$Im**, or glycols formed by the nucleophilic attack of the C5/C4 carbons on the carbonyl group of glycolaldehyde.

References

1. Vázquez-Salazar, A.; Tan, G.; Stockton, A.; Fani, R.; Becerra, A.; Lazcano, A. *Orig. Life Evol. Biosph.* **2016**, *doi:10.1007/s11084-016-9525-y*.
2. Oró, J.; Basile, B.; Cortes, S.; Shen, C.; Yamrom, T. *Orig. Life Evol. Biosph.* **1984**, *14*, 237.
3. Shen, C.; Yang, L.; Miller, S. L.; Oró, J. *Orig. Life Evol. Biosph.* **1987**, *17*, 295.
4. Grimmett, M. R.; Richards, E. L. *Aust. J. Chem.* **1964** *17*, 1379.
5. Ferris, J. P.; Narang, R. S.; Newton, T. A.; Rao, V. R. *J. Org. Chem.* **1979**, *44*, 1273.
6. Bigot, B.; Roux, D. *J. Org. Chem.* **1981**, *46*, 2872.
7. Ferris, J. P.; Orgel, L. E. *J. Am. Chem. Soc.* **1966**, *88*, 1074.
8. Ritson, D.; Sutherland, J. D. *Nature Chem.* **2012**, *4*, 895.
9. Bouregeois, S.; Diakite, D.; Perdereau, M. *React. Sol.* **1988**, *6*, 65.
10. Summers, D. P.; Chang, S. *Nature* **1993**, *365*, 630.
11. Sanchez, R. A.; Orgel, L. E. *J. Mol. Biol.* **1970**, *47*, 531.
12. Li, L.; Prywes, N.; Tam, C. P.; O'Flaherty, D. K.; Lelyveld, V. S.; Izgu, E. C.; Pal, A.; Szostak, J. W. *J. Am. Chem. Soc.* **2017**, *139*, 1810.