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

ABSTRACT: We have recently shown that 2-aminoimidazole is a superior nucleotide activating group for nonenzymatic RNA copying. Here we describe a prebiotic synthesis of 2-aminoimidazole that shares a common mechanistic pathway with that of 2-aminooxazole, a previously described key intermediate in prebiotic nucleotide synthesis. In the presence of glycolaldehyde, cyanamide, phosphate and ammonium ion, both 2-aminoimidazole and 2-aminooxazole are produced, with higher concentrations of ammonium ion and acidic pH favoring the former. Given a 1:1 mixture of 2-aminoimidazole and 2-aminooxazole, glyceraldehyde preferentially reacts and cyclizes with the latter, forming a mixture of pentose aminooxazolines, and leaving free 2-aminoimidazole available for nucleotide activation. The common synthetic origin of 2-aminoimidazole and 2-aminooxazole and their distinct reactivities are suggestive of a reaction network that could lead to both the synthesis of RNA monomers and to their subsequent chemical activation.

Imidazoles are thought to have played important roles in prebiotic chemistry prior to their current biological significance as components of the histidine residues in proteins, where they function as, for example, charge-relay agents in catalytic triads. One of the most important potential roles of imidazoles in prebiotic chemistry came to light through the efforts of Orgel and co-workers, who demonstrated that nucleoside 5'-phosphoro-imidazolides, and especially nucleotides activated with 2-methylimidazole, allow for nonenzymatic template-directed RNA synthesis yielding predominantly the canonical 3'-5' phosphodiester linkage. These phosphoroimidazolides are more reactive than nucleoside triphosphates as a result of their labile P=O bonds and their propensity to form a reactive imidazolium-bridged dinucleotide. In addition to prebiotic activation chemistry, imidazoles have also been suggested to play catalytic roles, assisting in the oligomerization of amino acids, phosphates and nucleotides, and a number of plausible mechanisms for the prebiotic synthesis of imidazoles have been reported (further discussion in Supporting Information).

Scheme 1. Common Prebiotic Synthetic Pathway for 2-Aminooxazole (2NH2Ox) and 2-Aminoimidazole (2NH2Im)\textsuperscript{a}

\textsuperscript{a}P\textsubscript{i}: inorganic phosphate.

Other than prebiotic chemistry, compounds containing imidazole motifs, including 2-aminoimidazole and 2-thiimidazole, are known for their medicinal properties. We became interested in finding a prebiotic synthesis of 2-aminoimidazole (2NH2Im), because we recently demonstrated that ribonucleoside 5'-monophosphates activated with 2NH2Im offer a 10–100-fold enhancement in the rate of nonenzymatic template-directed RNA primer extension compared to those activated with 2-methylimidazole, i.e., in this context 2NH2Im is the most effective leaving group known to date. Furthermore, the conjugate base of 2NH2Im is isosteric and isoelectronic with 2-aminooxazole—a key intermediate in the prebiotic synthesis of cytidine and uridine 2',3'-cyclic phosphates, as reported by Sutherland and co-workers. The prebiotic synthesis of 2NH2Ox is known, and begins with the addition of glycolaldehyde and cyanamide, followed by a series of steps that are facilitated by inorganic phosphate acting as both a pH buffer and general-base catalyst. In this pathway, cyanamide first undergoes addition to the carbonyl of glycolaldehyde. Next, an intramolecular general-base-catalyzed attack of the glycolaldehyde-derived hydroxyl on the cyanamide-derived nitrile carbon leads to a five-membered ring. General-base catalysis by phosphate also likely accelerates C=H deprotonation in the following dehydration step leading to the aromatic 2NH2Ox.
Given their structural similarities, we wondered whether 2NH2Im and 2NH2Ox could share a common prebiotic synthetic pathway (Scheme 1). Previously reported (nonprebiotic) syntheses of 2NH2Im make use of the dimethyl- or diethyl-acetal of α-aminoacetaldehyde,19−22 which presumably stabilizes the amino group by preventing enolization and subsequent exchange with water. We reasoned that in order to synthesize 2NH2Im, we would need to find a plausible route for the synthesis of α-aminoacetaldehyde. We hypothesized simply adding a sufficiently high concentration of an ammonium salt to the mixture of glycolaldehyde, cyanamide and inorganic phosphate, would lead to the formation of at least some of this intermediate at equilibrium through an Amadori-type enol-mediated exchange mechanism (Figure 1a, Scheme S1). Alternatively, exchange of the glycolaldehyde-derived hydroxyl to an amine could take place after the addition of cyanamide to glycolaldehyde.

To test this hypothesis, we carried out reactions with aqueous solutions of cyanamide and glycolaldehyde in the presence of 1 M phosphate at varying concentrations of NH4Cl at pH 7, 60 °C for 3 h. In the absence of NH4Cl, analysis of the reaction mixture by 1H NMR spectroscopy revealed that 2NH2Ox is produced nearly exclusively (Figures S1 and S2), as expected based on previous reports by Sutherland et al.17,18 The addition of 1 M NH4Cl (Figure 1b) resulted in the appearance of another resonance in the aromatic region of the proton NMR spectrum. Addition of an authentic standard of 2NH2Im confirmed that this resonance arises from the H4 and H5 protons of 2NH2Im. Increasing the concentration of NH4Cl to 5 M resulted in the almost exclusive formation of 2NH2Im. Quantification of the yield was carried out using a calibrated solution of S′-cytidine monophosphate (CMP), which was added directly to the NMR tube; we used its H5 and H6 resonances as standards for integration. After 1 h, the reaction was complete, and the yield was determined to be 15%. Although this yield is not as high as that previously reported18 for 2NH2Ox (>80%), it is important to note that no other major products could be observed in the 1H NMR spectrum.

Understanding how the ratio of 2NH2Im to 2NH2Ox varies under different pH regimes and ammonium ion concentrations is important for evaluating the reaction in the context of potential geochemical scenarios. We systematically examined the reaction at pH 4, 5.5, 7 and 8.5 with NH4Cl concentrations of 0, 0.5, 1, 2, 3, 4 and 5 M. All reactions were monitored by 1H NMR after 1, 2 and 3 h, and the ratios of 2NH2Im to 2NH2Ox were measured by resonance integration (Table S1). The ratios after 3 h are plotted in Figure 2. For all pH values, the ratio of 2NH2Im to 2NH2Ox increases with increasing NH4Cl, whereas mildly acidic pH also tends to increase this ratio. At least part of this effect of pH can be explained by the fact that increasing acidity does not tend to favor dehydration of the 4-hydroxy intermediate of 2NH2Ox, a minor species that can be detected under the acidic conditions tested. A maximum ratio of [2NH2Im]/[2NH2Ox] = 56 was obtained at a pH value of 5.5 with 5 M NH4Cl after 3 h. Analysis by Q-TOF LC-MS confirmed these results (Table S2, Figures S3 and S4 for further details). Similar yields for both 2NH2Ox and 2NH2Im were obtained when either 100 or 50 mM concentrations of the starting materials were used. We carried out a preparative scale reaction for the synthesis of 2NH2Im at a pH of ~5.3 employing 1 M NH4H2PO4 and 5 M NH4HCO3 and obtained a 41% isolated yield (Scheme S2). Given the general mechanistic features of the above synthetic reactions, we anticipated the prebiotic formation of imidazoles may be rather general. We set out to make 2-thioimidazole (Figure S5). Heating a 1 M solution of NH4SCN and glycolaldehyde at pH 4 in 4 M NH4Cl at 60 °C for 24 h led to the formation of 2-thioimidazole with no other major products observable by 1H NMR, although the yield was relatively low at 6.2%.

Knowing that 2NH2Ox and 2NH2Im can be made in the same reaction flask in approximately equimolar ratios at pH 7 and ~1 M NH4Cl, we asked whether these two compounds would display sufficiently different reactivities such that 2NH2Ox could be channeled toward nucleotide synthesis, whereas 2NH2Im would be preserved for later nucleoside 5′-phosphate activation. The next step on the pathway from
2NH₂Ox to the cytidine and uridine cyclic phosphates involves a cycloaddition reaction with glyceraldehyde, which generates a mixture of ribo- and arabino-furanosyl aminooxazolines as the major products. For 2NH₂Im to be preserved for subsequent nucleotide activation, it would have to be significantly less reactive with glyceraldehyde than 2NH₂Ox. To address this question, we reacted a 1:1 mixture of 2NH₂Ox and 2NH₂Im with glyceraldehyde at 40 °C in the presence of inorganic phosphate for 24 h while monitoring the reaction mixture by high-resolution LC-MS (Figure S6). After 24 h, nearly all the 2NH₂Ox had been depleted from the reaction mixture, whereas ~80% of the 2NH₂Im remained. Analysis by 1H NMR spectroscopy confirmed this result (Figure S7). We also detected lesser amounts of a product with an m/z of 174.09, a value that is consistent with the [M + H]+ mass of the analogous product but cyclized with 2NH₂Im. Although it would appear that the 2NH₂Im can also react in some fashion with glyceraldehyde, based on the initial rates measured from control experiments reacting 2NH₂Im and 2NH₂Ox with glyceraldehyde individually in separate solutions, we estimate that the reaction with 2NH₂Ox is about 1 order of magnitude faster. Indeed, the reaction of 2NH₂Im with glyceraldehyde is sufficiently slow such that the reaction stops after the consumption of only ~20% of 2NH₂Im. The likely reason for this effect is that the isomerization of glyceraldehyde to dihydroxycetone is a competing process occurring at a similar rate, preventing the reaction from going to a completion. We suspect that the slower reaction kinetics of 2NH₂Im compared to 2NH₂Ox are at least in part a result of the greater aromatic stability of 2NH₂Im, the greater nucleophilicity of 2NH₂Ox, or both. We also examined the case of a “one-pot” reaction, in which a solution of 1 M cyanamide, glycolaldehyde, glyceraldehyde, ammonium chloride and sodium phosphate at pH 7 was heated to 40 °C and monitored by LC-MS and NMR over time (Figures S8 and S9). These results provide additional evidence that the greater stability of 2NH₂Im allows it to persist and potentially accumulate for later activation chemistry.

Finally, we demonstrate that N-cyano-2-aminimidazole (2NH₂ImCN) is capable of activating CMP to furnish (Scheme 2) cytidine-5′-phosphoro-(2-aminimidazole) (2NH₂ImpC). Several reports in the literature suggest that N-cyanoimidazoles can serve as an activating agent for the formation of phosphodiester bonds. 2NH₂ImCN appeared attractive as a potentially prebiotic activating agent, because it represents one of the simplest possible extensions of the chemistry reported here, i.e., oxidative coupling with hydrogen cyanide. We formed 2NH₂ImpC by reacting 2NH₂Im with cyanogen bromide in acetonitrile at room temperature for 15 min (Figure S10). The reaction was concentrated to near dryness, at which point an aqueous solution of CMP with 10% D₂O was added and the pH adjusted between 5.5 and 6. After about 10 min at room temperature, 31P NMR spectroscopic analysis revealed ~20% conversion to 2NH₂ImpC (Figure S11 and S12). A maximum of about 40% conversion was obtained after 1.6 h. By addition of another freshly prepared batch of 2NH₂ImCN, a final conversion of ~75% was achieved. While the prebiotic synthesis of 2NH₂ImpC has not yet been achieved, this synthesis of 2NH₂ImpC serves as a proof of concept, highlighting the potential usefulness of N-cyanoimidazoles as prebiotic activating agents.

In summary, we have demonstrated a prebiotically plausible synthetic pathway for 2-aminimidazole that shares a common origin with the synthesis of 2-aminooxazole. Recently, Powner et al. showed 2-aminothiazole can be efficiently synthesized in a more straightforward manner. 2NH₂ImpC is capable of activating CMP to furnish (Scheme 2) cytidine-5′-phosphoro-(2-aminimidazole) (2NH₂ImpC).

Scheme 2. Synthesis of Cytidine-5′-phosphoro-(2-aminimidazole) Making Use of N-cyano-2-aminimidazole
similar fashion as $2NH_2Ox$ and $2NH_2Im$, starting from cyanamide and β-mercaptoacetaldehyde.\(^{25}\) 2-Aminothiazole forms stable crystalline aminals with aldehydes, but not with ketones, allowing for the concomitant accumulation and purification of reactive aldehydes, and the chemical selection of proteinogenic amino acids. Remarkably, all three compounds in the series 2-amino-oxazole/imidazole/thiazole seem to have important potential prebiotic roles. In the shared pathway for $2NH_2Ox$ and $2NH_2Im$ production, the relative yield of each species depends on the pH and ammonium chloride concentration. At neutral pH, this proposed pathway for the prebiotic synthesis of $2NH_2Im$ requires high concentrations of aqueous ammonia, on the order of 1 M to generate comparable amounts of $2NH_2Im$ and $2NH_2Ox$. There are several scenarios in which ammonium ions could have been generated in a concentrated form within an ancient aqueous reservoir. In one scenario proposed by Sutherland,\(^{26}\) cyanide produced in the atmosphere rains out and is captured by ferrous ions as ferrocyanide; salts of ferrocyanide then precipitate and accumulate over long periods of time. Subsequent thermal processing of deposits of magnesium ferrocyanide by magma or moistening would hydrolyze, releasing ammonia. Other pathways to ammonia sources have also been suggested (see SI).

Finally, the greater stability/slower reactivity of $2NH_2Im$ in comparison to $2NH_2Ox$ suggests that the former could accumulate over time even as the latter is continuously processed into intermediates on the path to nucleotide synthesis. Having a mechanism for the simultaneous production of $2NH_2Im$ and $2NH_2Ox$ suggests the tantalizing prospect of an ancient prebiotic reaction network that could have led to subsequent chemical activation of those nucleotides in a manner suitable for the prebiotic activation of nucleoside $5'$-monophosphates, could have arisen from such a network.

**REFERENCES**

1. Vázquez-Salazar, A.; Tan, G.; Stockton, A.; Fani, R.; Becerra, A.; Lazcano, A. Origins Life Evol. Biospheres 2016, DOI: 10.1007/s11084-016-9525-y.
2. Dodson, G.; Wlodawer, A. Trends Biochem. Sci. 1998, 23, 347.
3. Wu, T.; Orgel, L. E. J. Am. Chem. Soc. 1992, 114, 5496.
4. Blain, J. C.; Szostak, J. W. Annu. Rev. Biochem. 2014, 83, 615.
5. Fahrenbach, A. C. Pure Appl. Chem. 2015, 87, 205.
6. Izu, E. C.; Fahrenbach, A. C.; Zhang, N.; Li, L.; Zhang, W.; Larsen, A. T.; Blain, J. C.; Szostak, J. W. J. Am. Chem. Soc. 2015, 137, 6373.
7. Li, L.; Lelyveld, V. S.; Prywes, N.; Szostak, J. W. J. Am. Chem. Soc. 2016, 138, 3986.
8. Walton, T.; Szostak, J. W. J. Am. Chem. Soc. 2016, 138, 11996.
9. Oró, J.; Basile, B.; Cortes, S.; Shen, C.; Yamrom, T. Origins Life 1984, 14, 237.
10. Lohmann, R.; Orgel, L. E. Nature 1973, 244, 418.
11. Weber, A. L. J. Mol. Evol. 1981, 18, 24.
12. Keeffe, A. D.; Miller, S. L. Origins Life Evol. Biospheres 1996, 26, 15.
13. Pongs, O.; Ts’o, P. O. P. J. Am. Chem. Soc. 1971, 93, 5241.
14. Oró, J.; Stephen-Sherwood, E. Origins Life 1974, 5, 159.
15. Zhang, L.; Peng, X.-M.; Danu, G. L. V.; Geng, R.-X.; Zhou, C.-H. Med. Res. Rev. 2014, 34, 340.
16. Li, L.; Prywes, N.; Tam, C. P.; O’Flaherty, D. K.; Lelyveld, V. S.; Izu, E. C.; Pal, A.; Szostak, J. W. J. Am. Chem. Soc. 2017, 139, 1810.
17. Anastasi, C.; Crowe, M. A.; Powner, M. W.; Sutherland, J. D. Angew. Chem., Int. Ed. 2006, 45, 6176.
18. Powner, M. W.; Geidt, B.; Sutherland, J. D. Nature 2009, 459, 239.
19. Lawson, A. J. Chem. Soc. 1956, 307.
20. Lancini, G. C.; Lazzari, E. J. Het. Chem. 1966, 3, 152.
21. Weinmann, H.; Harre, M.; Koenig, K.; Merten, E.; Titlsm, U. Tetrahedron Lett. 2002, 43, 593.
22. Ferris, J. P.; Huang, C.-H.; Hagan, W. J., Jr. Nucleosides Nucleotides 1989, 8, 407.
23. Chen, H.; Du, F.; Chen, G.; Streckenbach, F.; Yasmeen, A.; Zhao, Y.; Tang. Z. Sci. Rep. 2014, 4, 4595.
24. McCallum, P. B. W.; Weavers, R. T.; Grimmett, M. R.; Blackman, A. G. Aust. J. Chem. 1999, 52, 159.
25. Islam, S.; Bužar, D.-K.; Powner, M. W. Nat. Chem. 2017, 9, 584.
26. Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. Nat. Chem. 2015, 7, 301.

**ACKNOWLEDGMENTS**

J.W.S. is an Investigator of the Howard Hughes Medical Institute. This work was supported in part by grants from the Simons Foundation to J.W.S. (290363) and from the NSF (CHE-1607034) to J.W.S. A.C.F. is supported by a Research Fellowship from the Earth-Life Science Institute at the Tokyo Institute of Technology. L.L. is a Life Sciences Research Foundation Fellow. Part of this work was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas “Hadean Bioscience”, Grant Number JP26106003.