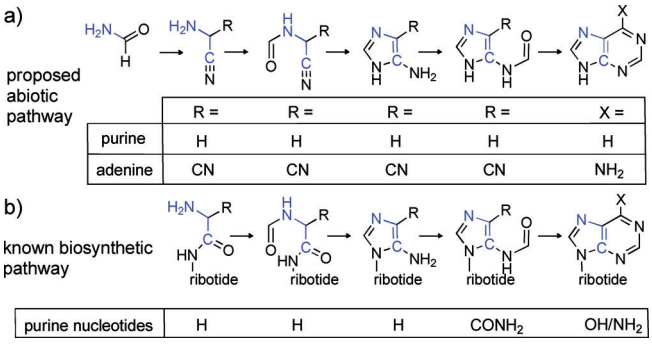
Prebiotic Synthesis

DOI: 10.1002/anie.201108907

A Unified Mechanism for Abiotic Adenine and Purine Synthesis in Formamide**

*Jeremy S. Hudson, Joseph F. Eberle, Raj H. Vachhani, Luther C. Rogers, James H. Wade, Ramanarayanan Krishnamurthy, and Greg Springsteen**

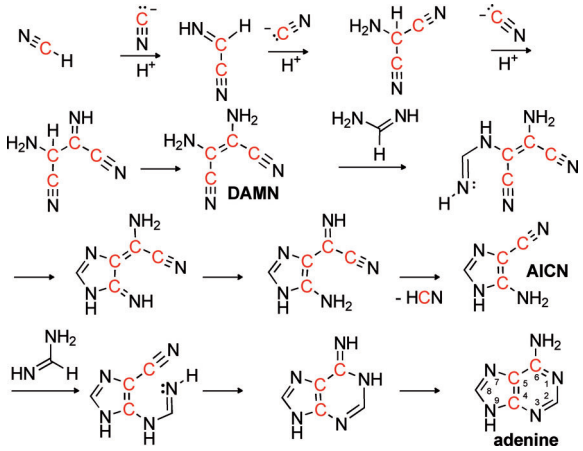
The synthesis of purines through dehydration and condensation of formamide is a potential abiotic source of nucleobases. It has been reported since the 1950s that heating formamide to near boiling generates purine in high yield (>70%).[1] The scope of this formamide condensation has been broadened by the identification of multiple biologically relevant products, including nucleobases and amino acid derivatives, from both neat and mineral-doped formamide.[2–4] Mechanistic pathways to purines have been proposed,[3–8] though there are significant variations between routes leading to related products. Herein, data is presented suggesting a common pathway for the abiotic syntheses of both purine and adenine from formamide; the proposed route is also highly reminiscent of the biosynthesis of purine nucleobases (Scheme 1). This is the first evidence suggesting that a glycine derivative is a critical intermediate in purine synthesis from formamide, and that this glycine backbone forms a scaffold for purine ring production in the same orientation as in purine ring biosynthesis. Parallels between nucleobase biosynthesis and plausible abiotic syntheses may allude to the origins of this metabolic pathway.[9]



	R =	R =	R =	R =	X =
purine	H	H	H	H	H
adenine	CN	CN	CN	CN	NH_2
purine nucleotides	H	H	H	$CONH_2$	OH/NH_2

Scheme 1. A glycine scaffold (blue) forms the ring juncture in the proposed abiotic pathway (a), in a manner reminiscent of the biosynthesis of purine nucleotides (b).

Abiotic nucleobases and their analogues have been identified in meteorites,[10] spark discharge experiments,[11] and hydrogen cyanide (HCN) condensation reactions.[12–15] Typical laboratory procedures call for one to 15 M HCN in aqueous ammonia at 25–70 °C.[12–14] Analyses of the non-photochemical, mechanistic route from HCN to adenine have been conducted by Oró, Orgel, and Ferris.[12–15] Briefly, diaminomaleonitrile (DAMN), a relatively stable HCN tetramer, has been identified as an important intermediate. It provides the ring junction scaffold that subsequently closes to an imidazole and then a bicyclic purine, through reaction with cyanide-derived formamidine (Scheme 2). In aqueous



Scheme 2. Literature-proposed route of non-photochemical adenine formation from HCN.[13–15] Red labels track cyanide carbon atoms. Formamidine carbon atoms are unlabelled, though formamidine may be derived by the amination of HCN. AICN = 5-aminoimidazole-4-cyanide

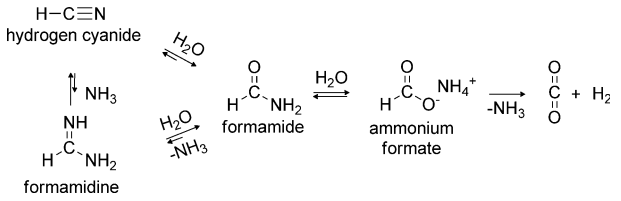
ammonia, hydrogen cyanide is partially aminated to formamidine, and hydrated to formamide and ammonium formate (Scheme 3). Reactions that replace the aqueous ammonia with formamide are also consistent with the DAMN mechanism; adenine carbons C^4, C^5, and C^6 are derived from HCN, while formamide provides C^2 and C^8 (Scheme 2).[5]

The synthesis of nucleobases in neat formamide has also been investigated.[2–7] The abiotic concentration of formamide, arising from dilute aqueous mixtures by evaporation, is more plausible than that of HCN. Formamide's vapor pressure is lower than that of both HCN and water, with

[*] J. S. Hudson, J. F. Eberle, R. H. Vachhani, L. C. Rogers, J. H. Wade, Dr. G. Springsteen
Department of Chemistry, Furman University
3300 Poinsett Hwy, Greenville, SC 29613 (USA)
E-mail: greg.springsteen@furman.edu
Prof. Dr. R. Krishnamurthy
Department of Chemistry, The Scripps Research Institute
10550 North Torrey Pines Rd., La Jolla, CA 92037 (USA)

[**] This work was jointly supported by NSF and the NASA Astrobiology Program under the NSF Center for Chemical Evolution, CHE1004570.

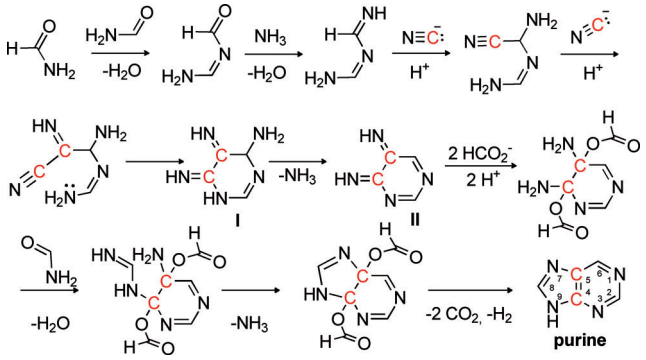
Supporting information for this article (experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201108907.



Scheme 3. Hydration and amination of hydrogen cyanide.

a wide liquid range at atmospheric pressure (2–210 °C). Dehydration of formamide to cyanide, though slow,[16] yields the carbon nucleophile necessary for the synthesis of complex organic compounds.

Purine is the most abundant nucleobase derivative produced by heating neat formamide.[2] Experimental and computational data support a pyrimidine-first mechanism in reactions starting from equimolar concentrations of HCN and formamide (Scheme 4).[7] This pathway bears little resemblance to the DAMN-based adenine synthesis from HCN and

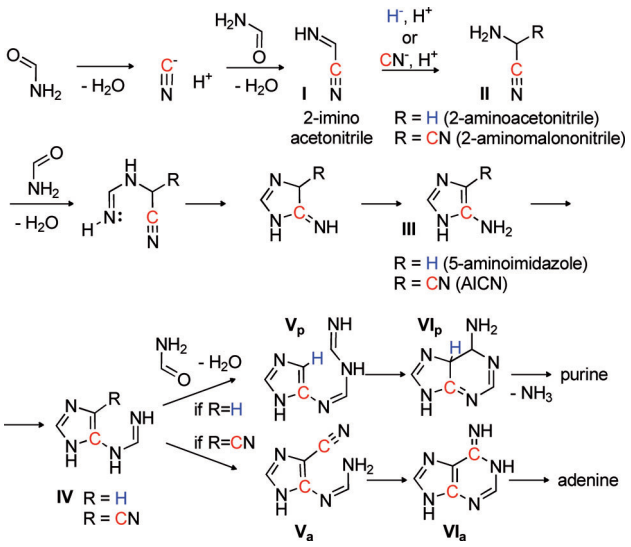


Scheme 4. Literature proposed route of purine formation from formamide.[7] Red labels track cyanide carbon atoms.

is complicated by the requirement for an in situ reduction. Pathways from formamide to adenine have been theorized to progress though pyrimidine[3] or DAMN[4] intermediates, though mechanistic experiments have not been conducted.

The mechanistic pathway presented herein begins with the dehydration of formamide to cyanide (Scheme 5).[16] Subsequent formylation and dehydration (formiminylation) generates 2-iminoacetonitrile (**I**, Scheme 5), a strong electrophile capable of progressing to either purine or adenine.

In the pathway to purine, a Leuckart reduction,[17] using hydride generated from formate, produces 2-aminoacetonitrile, a dehydrated glycine equivalent (Scheme 5, **II**, R = H). Formiminylation of 2-aminoacetonitrile, followed by ring closing and tautomerization, generates 5-aminoimidazole (**III**, R = H). In our hands, glycine, 2-aminoacetonitrile (**II**), and 5-aminoimidazole (**III**)[18] produced purine (in 2%, 10%, and 31% yields, respectively), when heated in 15 mol% ammonium formate in formamide at 150 °C for 18 h, our typical reaction conditions (Supporting Information, Figure S1). Subsequent formiminylation of **III** may occur on either the imidazole carbon or the exocyclic nitrogen atom.[19]



Scheme 5. Proposed pathway for the formation of purine and adenine in formamide. Red labels track cyanide-derived carbon atoms, blue label tracks formate-derived hydride.

Elucidation of the correct regioselectivity was achieved by isolating intermediate **IV** (R = H), as its hydrolyzed 5-*N*-formyl amide, in a reaction from 2-aminoacetonitrile at reduced temperature (100 °C). The purified product was analyzed by mass spectrometry (*m*/*z* 112.31 found, *m*/*z* 112.05 expected). The ^{1}H NMR spectrum displays signals for two imidazole hydrogen atoms at $\delta = 7.26$ and 7.75 ppm, consistent with N- rather than C-formiminylation (Supporting Information, Figures S2–S4). Subsequent intermediates were not isolated, suggesting purine production was facile after the formiminylation of **IV**. A ring-closing addition and tautomerization of $\mathbf{V_p}$ to $\mathbf{VI_p}$ (p indicates intermediate on the pathway to purine), followed by an aromatizing elimination of ammonia, completes the pathway to purine.

Competing with hydride for 2-iminoacetonitrile (**I**) is an additional cyanide nucleophile, which reacts to generate 2-aminomalononitrile (**II**, R = CN) on the pathway to adenine. The formiminylation of 2-aminomalononitrile, followed by ring closing and tautomerization, produces 5-aminoimidazole-4-cyanide (AICN, **III**, R = CN). The progression of AICN to adenine proceeds by formiminylation and tautomerization to $\mathbf{V_a}$, ring closing to $\mathbf{VI_a}$, and a final tautomerization to adenine. In our hands, 25 mg of aminomalononitrile produced adenine in 17% yield under typical reaction conditions. Hydrated isomers of **III** and **IV** (R = CN) have previously been identified in mineral-doped formamide condensation reactions.[2c]

Consistent with the proposed mechanisms are product distributions resulting from varied starting concentrations of ammonium formate and cyanide. The pathways to purine and adenine compete for 2-iminoacetonitrile, thus an increased concentration of hydride-generating formate favors the reduction pathway to purine. A reaction of 5.0 mg of KCN in 1.0 mL of formamide favors adenine production over purine production by 5.5:1 (Table 1). However, the same reaction, but with 15 mol% formate favors purine by over 6:1.

Table 1: Adenine and purine production as a function of [formate].[a]

HCO_2NH_4 [mol%]	Adenine [μg]	Purine [μg]	Adenine:Purine
0	220	40	5.5:1
1	150	80	1.9:1
2	240	190	1.3:1
5	210	330	1:1.6
10	90	290	1:3.2
15	40	250	1:6.3

[a] 5.0 mg of KCN in 1.0 mL of 0–15 mol% ammonium formate in formamide heated for 2 h at 165 °C.

Correspondingly, increased cyanide favors adenine formation (Table 2); elevating the starting amount of KCN from 2–30 mg raises adenine yield by 140-fold while decreasing purine yield. The dependence of adenine formation on initial cyanide concentration is consistent with a reaction that is second-order in cyanide concentration, as expected from the proposed mechanism.

Table 2: Adenine and purine production as a function of [cyanide].[a]

KCN [mg]	Adenine [μg]	Purine [μg]	Adenine:Purine
0	< 10	70	
2	10	260	1:26
5	40	250	1:6.3
9	110	230	1:2.1
14	360	230	1.6:1
30	1400	190	7.4:1

[a] 0–30.0 mg KCN in 1.0 mL of 15 mol% ammonium formate in formamide heated for 2 h at 165 °C.

Cyanide is the source of a single carbon in the proposed mechanism to purine (red-label, Scheme 5); the previously described pyrimidine-based mechanism consumes two cyanide carbons (red-label, Scheme 4). Evidence supporting the proposed mechanistic pathway in formamide (with low concentrations of cyanide) was obtained by heating 14 mg of $K^{13}CN$ in 1.0 mL of 15 mol% ammonium formate in formamide at 165 °C for 2 h. The reaction produced a 3:2 ratio of adenine:purine, which were separated by reverse phase-HPLC (RP-HPLC) and analyzed by NMR spectroscopy and high-resolution mass spectrometry (MS). The m/z of the purine product was measured at 122.0555 (m/z 122.0548 predicted for mono-labeled purine, Figure S6). The ^1H NMR spectrum of the purine fraction displays three doublets , each arising from a single ^1H-^{13}C coupling, consistent with a mono-labeled product and the proposed mechanism (Figure 1 a). The ^{13}C NMR spectrum has only one enhanced signal ($\delta = 157.4$ ppm), consistent with isotopic labeling at C^4 (Supporting Information, Figure S7). The production of di-labeled purine, as reported by Yamada et al.,[7a] increases with initial cyanide concentration, though the mono-labeled species predominates even at 100 mg/ml of $K^{13}CN$ (100:17 by selective ion count, Supporting Information, Figure S6, Table S1). An imidazole-first mechanism may also explain the di-labeled purine detected at high $K^{13}CN$ concentrations (Supporting Information, Scheme S1).

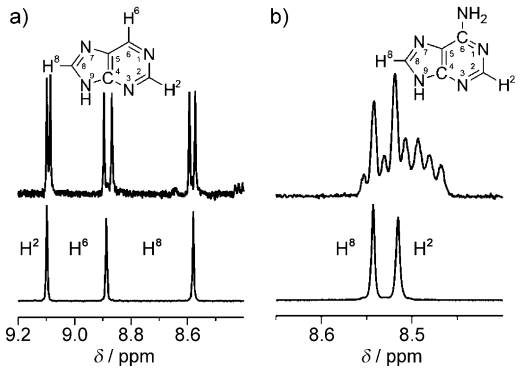


Figure 1. ^1H NMR spectra of a) purine and b) adenine, synthesized from ^{13}C-labeled KCN (top), and unlabeled standards (bottom). Spectra recorded in [D_6]DMSO (purine) and [D_6]DMSO/DCl (adenine).

Reactions in formamide that begin with high initial concentrations of ^{13}C-labeled HCN generate tri-labeled adenine,[5] implicating a DAMN mechanism (Scheme 2). However, in our experiments, which contain low starting cyanide concentrations ($K^{13}CN$) as a better model of 150–180 °C formamide, the adenine fraction is primarily di-isotopically labeled, consistent with the proposed mechanism (Scheme 5).MS analysis of the product reveals m/z 138.0685 (m/z 138.0690 predicted for di-labeled adenine, Supporting Information, Figure S9). In the ^1H NMR spectrum, two aromatic adenine protons are each split by two ^{13}C atoms, resulting in a pair of doublet of doublets (8 peaks, Figure 1 b). The ^{13}C NMR spectrum has two isotopically enhanced peaks consistent with C^4 and C^6 at $\delta = 149.3$ and 150.6 ppm, respectively (Supporting Information, Figure S9, S10). The DAMN reaction pathway is observable as a minor secondary route; labeled carbon at C^5 ($\delta = 113.9$ ppm) split by C^4 ($J =$ 256 hz) and C^6 ($J = 304$ hz) is consistent with the tri-labeled adenine expected for this pathway (Supporting Information, Figure S9).

In summary, we have described a unified mechanistic pathway to adenine and purine in formamide. The order and chemoselectivity of this route is reminiscent of purine biosynthesis. Observations of such similarities may help elucidate the evolution of early biotic metabolism.

Received: December 17, 2011
Revised: February 9, 2012
Published online: April 5, 2012

Keywords: cyanides · formamide · nucleobases · prebiotic chemistry · purines

[1] a) H. Yamada, T. Okamoto, *Chem. Pharm. Bull.* **1972**, *20*, 623–624; b) H. Bredereck, F. Effenberger, G. Rainer, H. P. Schosser, *Justus Liebigs Ann. Chem.* **1962**, *659*, 133–138; c) H. Bredereck, H. Waldmann, DE 941289, **1956**, CAplus **1958**:82968.

[2] a) R. Saladino, C. Crestini, G. Costanzo, R. Negri, E. Di Mauro, *Bioorg. Med. Chem.* **2001**, *9*, 1249–1253; b) R. Saladino, U. Ciamecchini, C. Crestini, G. Costanzo, R. Negri, E. Di Mauro, *ChemBioChem* **2003**, *4*, 514–521; c) R. Saladino, C. Crestini,, U. Ciamecchini, F. Ciciriello, G. Costanzo, E. Di Mauro, *Chem-*

BioChem **2004**, *5*, 1558 – 1566; d) R. Saladino, V. Neri, C. Crestini, G. Costanzo, M. Graciotti, E. Di Mauro, *J. Am. Chem. Soc.* **2008**, *130*, 15512 – 15518.
[3] a) R. Saladino, C. Crestini, G. Costanzo, E. Di Mauro, *Curr. Org. Chem.* **2004**, *8*, 1425 – 1433; b) R. Saladino, C. Crestini, F. Ciciriello, G. Costanzo, E. Di Mauro, *Chem. Biodiversity* **2007**, *4*, 694 – 720.
[4] H. L. Barks, R. Buckley, G. A. Grieves, E. Di Mauro, N. V. Hud, T. M. Orlando, *ChemBioChem* **2010**, *11*, 1240 – 1243.
[5] H. Yamada, M. Hirobe, K. Higashiyama, H. Takahashi, K. T. Suzuki, *J. Am. Chem. Soc.* **1978**, *100*, 4617 – 4618.
[6] M. Ochiai, R. Marumoto, S. Kobayashi, H. Shimazu, K. Morita, *Tetrahedron* **1968**, *24*, 5731 – 5737.
[7] a) H. Yamada, M. Hirobe, K. Higashiyama, H. Takahashi, K. T. Suzuki, *Tetrahedron Lett.* **1978**, *19*, 4039 – 4042; b) H. Yamada, M. Hirobe, T. Okamoto, *Yakugaku Zasshi* **1980**, *100*, 489 – 492; c) J. E. Šponer, A. Mládek, J. Šponer, M. Fuentes-Cabrera, *J. Phys. Chem. A* **2011**, *116*, 720 – 726.
[8] R. F. Shuman, W. E. Shearin, R. J. Tull, *J. Org. Chem.* **1979**, *44*, 4532 – 4536.
[9] a) C. Katsigiannis, A. Mar, J. Oró, *Origins Life Evol. Biosphere* **1986**, *16*, 297 – 298; b) I. M. Lagoja, P. Herdewijn, *Chem. Biodiversity* **2004**, *1*, 106 – 111.
[10] a) Z. Peeters, O. Botta, S. B. Charnley, R. Ruiterkamp, P. Ehrenfreund, *Astrophys. J. Lett.* **2003**, *593*, L129 – L132; b) M. P. Callahan, K. E. Smith, H. J. Cleaves, J. Ruzicka, J. C. Stern, D. P. Glavin, C. H. House, J. P. Dworkin, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13995 – 13998.
[11] S. Miyakawa, H. Yamanashi, K. Kobayashi, H. J. Cleaves, S. L. Miller, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14628 – 14631.
[12] a) J. Oró, *Biochem. Biophys. Res. Commun.* **1960**, *2*, 407 – 412; b) J. Oró, *Nature* **1961**, *191*, 1193 – 1194.
[13] a) R. A. Sanchez, J. P. Ferris, L. E. Orgel, *J. Mol. Biol.* **1967**, *30*, 223 – 253; b) J. P. Ferris, L. Orgel, *J. Am. Chem. Soc.* **1966**, *88*, 1074.
[14] a) J. Oró, A. P. Kimball, *Arch. Biochem. Biophys.* **1961**, *94*, 217 – 227; b) R. A. Sanchez, J. P. Ferris, L. E. Orgel, *J. Mol. Biol.* **1968**, *38*, 121 – 128.
[15] a) J. P. Ferris, L. E. Orgel, *J. Am. Chem. Soc.* **1965**, *87*, 4976 – 4977; b) J. P. Ferris, L. E. Orgel, *J. Am. Chem. Soc.* **1966**, *88*, 3829 – 3831.
[16] a) F. Cataldo, G. Patane, G. Compagnini, *J. Macromol. Sci. Part A* **2009**, *46*, 1039 – 1048; b) V. S. Nguyen, H. L. Abbott, M. M. Dawley, T. M. Orlando, J. Leszczynski, M. T. Nguyen, *J. Phys. Chem. A* **2011**, *115*, 841 – 851.
[17] M. L. Moore in *Organic Reactions*, *Vol. 5* (Eds.: R. Adams, W. E. Bachmann, A. H. Blatt, L. F. Fieser, J. R. Johnson), Wiley, New York, **1949**, pp. 301 – 330.
[18] Z. Wang, H. K. Huynh, B. Han, R. Krishnamurthy, A. Eschenmoser, *Org. Lett.* **2003**, *5*, 2067 – 2070.
[19] A. H. M. Al-Shaar, R. K. Chambers, D. W. Gilmour, D. J. Lythgoe, I. McClenaghan, C. A. Ramsden, *J. Chem. Soc. Perkin Trans. 1* **1992**, 2789 – 2811.

