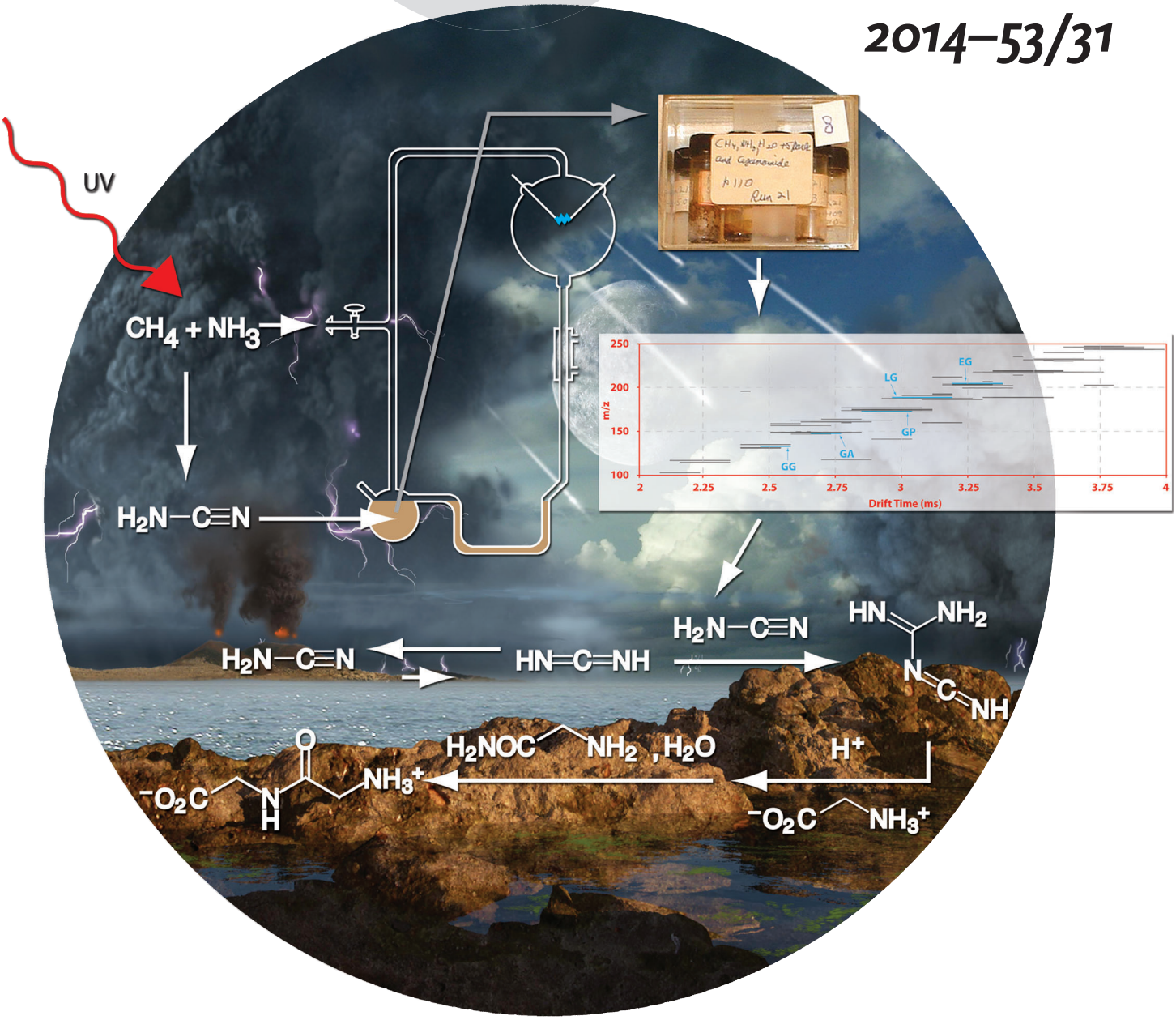
A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition Chemie

GDCh

www.angewandte.org

2014–53/31



Stanley Miller's cyanamide experiment ...

... went unreported for over 50 years, but was recently explored to study cyanamide-mediated biomolecule polymerization under early Earth conditions. In their Communication on page 8132 ff., F. M. Fernández, J. L. Bada et al. show that the dimerization of cyanamide in the presence of amino acids and intermediates in the Strecker synthesis of amino acids yields significant levels of dipeptides, which provides evidence that cyanamide enhances polymerization under simulated prebiotic environments. (Background image provided by Ron Miller.)

WILEY-VCH

Prebiotic Chemistry Hot Paper

DOI: 10.1002/anie.201403683

A Plausible Simultaneous Synthesis of Amino Acids and Simple Peptides on the Primordial Earth**

Eric T. Parker, Manshui Zhou, Aaron S. Burton, Daniel P. Glavin, Jason P. Dworkin, Ramanarayanan Krishnamurthy, Facundo M. Fernández, and Jeffrey L. Bada**

Dedicated to Stanley L. Miller and Joan Oró

Abstract: *Following his seminal work in 1953, Stanley Miller conducted an experiment in 1958 to study the polymerization of amino acids under simulated early Earth conditions. In the experiment, Miller sparked a gas mixture of CH_4, NH_3, and H_2O, while intermittently adding the plausible prebiotic condensing reagent cyanamide. For unknown reasons, an analysis of the samples was not reported. We analyzed the archived samples for amino acids, dipeptides, and diketopiperazines by liquid chromatography, ion mobility spectrometry, and mass spectrometry. A dozen amino acids, 10 glycine-containing dipeptides, and 3 glycine-containing diketopiperazines were detected. Miller's experiment was repeated and similar polymerization products were observed. Aqueous heating experiments indicate that Strecker synthesis intermediates play a key role in facilitating polymerization. These results highlight the potential importance of condensing reagents in generating diversity within the prebiotic chemical inventory.*

Stanley Miller published the synthesis of amino acids by sparking a gas mixture of methane, ammonia, water, and hydrogen,[1] which were considered in the early 1950s to be representative of the early Earth's atmosphere.[2] Today, however, a weakly reducing or neutral primitive terrestrial atmosphere comprised of major constituents such as CO_2, N_2,[3] CO, and H_2O, with minor components, including reduced gases such as H_2, H_2S, and CH_4,[4] is favored to a strongly reducing gas mixture. Although reducing atmospheric conditions may have been unlikely on a global scale on the early Earth, they might have been present on smaller scales[5] that could have been important locales capable of fostering a suite of very powerful prebiotic chemical reactions to produce large quantities of molecules important for life.[5a,b] Laboratory studies have shown that, even under neutral conditions, amino acid synthesis is efficient.[4]

A combination of Miller's pioneering 1953 experiment[1] and the subsequent findings of extraterrestrial organic compounds in meteorites[6] indicates that the synthesis of prebiotic organic compounds thought to be necessary for the origin of life is a robust process, both on the primitive Earth and on other planetary bodies.[7] However, the transition from simple molecules, such as amino acids, to more complex ones, such as peptides, has proven challenging under plausible primordial conditions. Although the syntheses of peptides by hydrothermal vents and comet impact have been reported, questions remain about their plausibility under prebiotic geochemical conditions.[4] In addition, concentrated salts, clays, and Cu^{2+} ions have been suggested as being important amino acid condensation reagents,[8] although these have not been demonstrated to be effective polymerization agents under the natural geochemical environments that may have existed on the early Earth. For example, Cu^{2+} ions in the primitive oceans would have been in the form of Cu^{+} and its concentration would have been very low because of the presence of HS^{-}.[9] Additionally, other reduced metal ions,

[*] E. T. Parker, Dr. M. Zhou, Prof. Dr. F. M. Fernández
School of Chemistry and Biochemistry
Georgia Institute of Technology
901 Atlantic Drive NW, Atlanta, GA 30332 (USA)
E-mail: facundo.fernandez@chemistry.gatech.edu
Dr. A. S. Burton
Astromaterials Research and Exploration Science Directorate, NASA Johnson Space Center, Houston, TX 77058 (USA)
Dr. D. P. Glavin, Dr. J. P. Dworkin
Solar System Exploration Division
NASA Goddard Space Flight Center, Greenbelt, MD 20771 (USA)
Prof. Dr. R. Krishnamurthy
Department of Chemistry, The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Prof. Dr. J. L. Bada
Geophysical Research Division, Scripps Institution of Oceanography, University of California, San Diego
8615 Kennel Way, La Jolla, CA 92093 (USA)
E-mail: jbada@ucsd.edu

[**] This research was supported by the Center for Chemical Evolution at the Georgia Institute of Technology, jointly supported by the NSF and the NASA Astrobiology Program (NSF CHE-1004570). E.T.P. acknowledges financial support from the Marine Biology Laboratory's NASA Planetary Biology Internship Program. A.S.B. acknowledges support from the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by Oak Ridge Associated Universities through a contract with NASA. J.P.D. and D.P.G. acknowledge the Goddard Center for Astrobiology and the NASA Astrobiology Institute. We are appreciative of the Mandeville Special Collections at the Geisel Library on the campus of the University of California, San Diego for archiving Miller's original laboratory notebooks and providing assistance with retrieving them.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201403683.

such as Fe^{2+}, may have been present[10] and could have played a role in shaping prebiotic environments and the chemical reactions that may have occurred therein. Other potential prebiotic polymerization agents such as carbonyl sulfide (COS) have been proposed,[11] although the overall efficiency with respect to the variety of amino acids that can undergo polymerization with this reagent has not been explored.

In carbonaceous meteorites, the timescale for the production of amino acids is typically estimated to be between 10^3 and 10^6 years,[12] but could be as short as 1–10 years.[13] The prebiotic chemistry that took place on the meteorite parent bodies during the aqueous alteration phase is considered to have produced mainly simple monomeric compounds[7a] and complex, poorly characterized, polymers.[14] Only very low, trace, quantities of glycine dipeptide and its diketopiperazine (DKP) have been detected in a few cases.[15]

Recently, archived stored portions of the solutions from experiments Miller carried out in 1958 were found.[5b] Included in this set were labeled vials from an experiment in which reduced gases (methane and ammonia) were subjected to a spark discharge for about 7 days, and over the course of the experiment, cyanamide was intermittently added to the aqueous phase. These archived cyanamide samples were part of a large collection of samples that Miller saved from a number of his experiments in the 1950s.[7a] For unknown reasons, Miller never performed a chemical analysis of the products of the cyanamide experiment and others conducted in 1958.[5b]

Although it has not yet been demonstrated that cyanamide can be formed within electric discharge experiments, the production of cyanamide in plausible prebiotic conditions comprised of CH_4, NH_3, H_2O, and UV light was reported nearly 50 years ago, and was proposed to be a possible prebiotic condensing reagent.[16] Preliminary experiments supported this scenario,[17] although the reaction seemed to be most favorable at acidic pH values lower than the pK_1 value of the amino acid.[18] This finding suggested that the reactive amino acid species is H_3N^+-C(RR′)-COO^- (where R and R′ represent the amino acid side chains). It has also been proposed that cyanamide can activate *N*-acyl-α-amino acids to form a 5(4*H*)-oxazolone intermediate that can help facilitate the coupling of sterically hindered α-amino acids.[19] In addition, cyanamide has been suggested to be influential in other important prebiotic reactions, such as the synthesis of activated pyrimidine ribonucleotides[20] and 2′-deoxynucleotides.[21]

Miller never carried out a detailed analysis of his 1958 cyanamide experiment, but he did measure the absorption at 280 nm when he collected various fractions during chromatographic separation of the discharge solution from the cyanamide experiment and found absorption in several samples where peptides were expected to elute.[22] We have now analyzed the 1958 cyanamide spark discharge residues to ascertain if both amino acids and simple peptides had actually been synthesized simultaneously in this prebiotic simulation experiment. Amino acids were analyzed using high-performance liquid chromatography with fluorescence detection and triple quadrupole mass spectrometry. Dipeptides and DKPs were identified and quantified using ultraperformance liquid chromatography coupled to quadrupole-traveling wave ion mobility spectrometry/time of flight mass spectrometry (for further details on the analytical tools used in this study, please see the Supporting Information).

The analysis of Miller's archived cyanamide experiment samples resulted in the detection of 12 amino acids, 10 glycine-containing dipeptides, and 3 glycine-containing DKPs (Table 1). The amino acids produced by the cyanamide experiment were synthesized in relatively high yields, and present in similar relative abundances compared to those detected in Miller's classic and volcanic,[5a] as well as hydrogen sulfide containing[5b] spark-discharge experiments (Figure 1). Major amino acids with stereogenic centers (e.g. aspartic and glutamic acids, serine, alanine, and isovaline) were racemic (D/L ≈ 1:1) within error limits (10%), thus

Table 1: Amino acids, dipeptides, and DKPs that were both identified and quantified in this study.

Amino acids[a]	Dipeptides[b]	DKPs
glycine	Gly-Gly	cyclo(Gly-Gly)
alanine	Gly-Ala	cyclo(Gly-Pro)
β-alanine	Gly-Thr	cyclo(Leu-Gly)
serine	Gly-Pro	
α-aminobutyric acid	Pro-Gly	
β-aminobutyric acid	Gly-Val	
γ-aminobutyric acid	Val-Gly	
aspartic acid	Gly-Glu	
glutamic acid	Glu-Gly	
valine	Leu-Gly	
isovaline		
isoleucine		

[a] Additional amino acids were tentatively identified, but were not quantified and, therefore, not included here, but are listed in the Supporting Information. [b] Additional dipeptides, as well as higher order peptides, such as the tripeptides Pro-Pro-Gly and Asp-Asp-Gly, were tentatively identified within the archived samples and are also shown in the Supporting Information. These initial identifications indicate that the formation of tri- and higher peptides in prebiotic simulation experiments warrants further investigation.

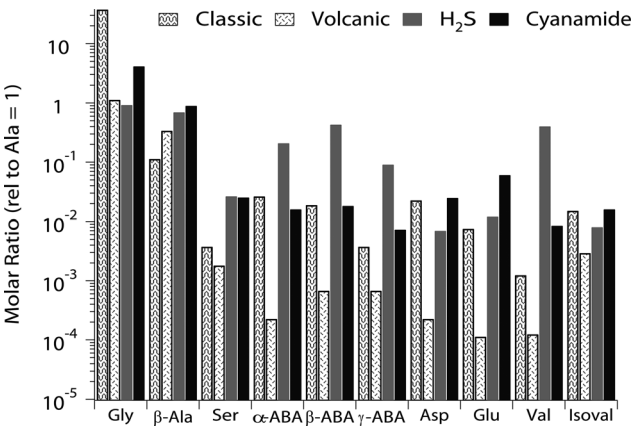


Figure 1. Molar ratios (relative to D + L-Ala = 1) of major amino acids in unhydrolyzed samples from the classic, volcanic, H_2S, and cyanamide spark-discharge experiments.

indicating their abundances were minimally influenced by contamination with terrestrial L-amino acids during sample storage and processing.

Glycine-containing dipeptides and DKPs were targeted for analysis. Glycine is the simplest amino acid and is one of the most abundant amino acids formed in spark-discharge experiments. Therefore, many peptides present in the samples reported here should contain glycine. Multiple analysis workflows (see the Supporting Information) were used to confirm the identity and quantity of the dipeptides and DKPs (Figure 2). The ratio of amino acids to dipeptides in the

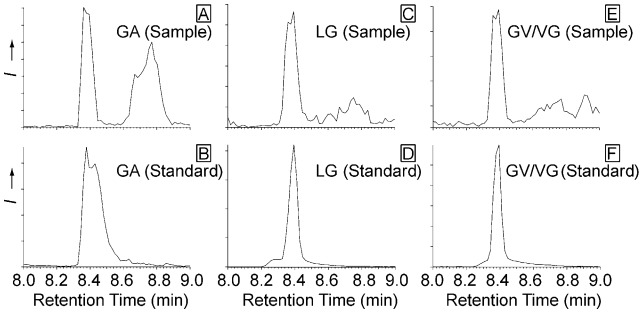


Figure 2. An example of how glycine-containing dipeptides were initially detected. Accurate-mass ultraperformance liquid chromatography (UPLC)/mass spectrometry analysis screened for specific dipeptide $[M+H]^+$ ions to provide a preliminary detection of the dipeptides of interest based on retention time and high-resolution exact mass, compared to standards. After this step, a more rigorous analytical approach was taken for unambiguous identification and quantification of the targeted dipeptides (see the Supporting Information). Extracted UPLC mass chromatograms, corresponding to the 8–9 min retention time window are displayed for specific dipeptides detected in Miller's sample fraction 57–67, and for standard traces. Chromatograms (A) and (B) were extracted by screening the total ion chromatograms for m/z 147.0770, C) and D) correspond to m/z 189.1239, and E) and F) correspond to m/z 175.1083. GA = glycylalanine, LG = leucylglycine, and GV/VG = glycylvaline/valylglycine.

cyanamide samples was calculated to be approximately 1000:1–1000:10, which agrees well with experimental data that indicates that the amino acid to dipeptide ratio is approximately 1000:1 under equilibrium conditions.[23] Furthermore, experimental data suggest that, at equilibrium, the dipeptide to DKP ratio should be on the order of 1:10,[24] and this ratio was determined to be 1:10–1:20.5 in the samples studied here. The cyclic nature of the DKP is responsible for its higher thermodynamic stability, and thus, greater abundance than the linear dipeptide at equilibrium.[24]

The presence of dipeptides in the archived samples was further confirmed by performing an acid hydrolysis on a portion of each sample,[25] analyzing the hydrolyzed fractions, and verifying that the peptide bonds had been cleaved to yield their amino acid residues. Additionally, identical dipeptide and DKP analyses, as reported for Miller's cyanamide samples, were carried out simultaneously on electric-discharge samples from Miller's 1958 hydrogen sulfide experiment,[5b] which did not incorporate a condensing reagent. Peptides were undetectable in the H_2S samples, thus providing added evidence to suggest that the presence of a condensing reagent facilitates the polymerization of amino acids.

In addition to investigating the archived cyanamide samples with modern analytical techniques, Miller's cyanamide experiment was repeated to generate fresh samples for further study. The analysis of the aqueous solution from the repeated experiment was compared to that of the original samples. The repeated experiment resulted in polymerization products, including dipeptides similar to those detected in the original 1958 cyanamide samples. These findings help corroborate the results obtained from the archived samples in suggesting that cyanamide can induce peptide formation under such a mimicked primitive Earth environment (the experimental and analytical details of this work are provided in the Supporting Information).

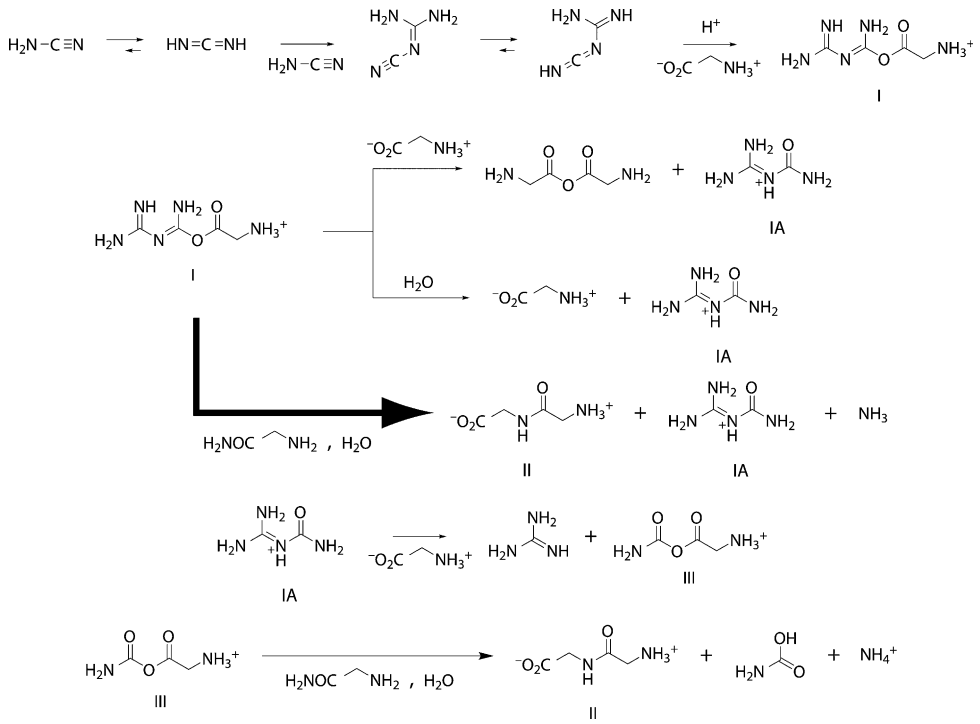
The formation of dipeptides in a mildly basic medium (pH 8–10) created by ammonia in the spark-discharge experiment contrasts with previous reports that indicate that acidic conditions are necessary to promote cyanamide-mediated polymerization of amino acids. As noted previously, in acid solutions with pH values less than the pK_1 value of the amino acid,[18] the reacting amino acid species would be H_3N^+-C(RR′)-COO^-. As the pK_a value of the COOH group is 2–2.5, the concentration of this reactive species decreases as the pH increases above pH ≈ 3, and the abundance of the protonated carboxylic acid is thus expected to be negligible at the pH value of the spark-discharge experiments. This suggests that perhaps one or more components intrinsic to the spark-discharge experiment may be responsible for facilitating the observed amino acid polymerization. Possible candidates include the amino acid amides and nitriles, both of which are intermediates in the Strecker reaction involved in amino acid synthesis.[7a,26]

Heating experiments on aqueous solutions were carried out to evaluate how dipeptide synthesis could proceed under mildly basic conditions. Solutions containing only amino acids in the presence of cyanamide or its dimer, dicyandiamide (2-cyanoguanidine), were prepared at pH 1–2, pH 6–7, pH 9–10, and pH 12–13 and heated at 50 °C. Although dicyandiamide was not directly introduced into the discharge apparatus, its potential as a condensing reagent was evaluated because cyanamide is known to dimerize readily in basic solutions[27] and because dicyandiamide is also a proposed prebiotic condensing reagent.[28] Analyses of the heated solutions at various pH values confirmed that dipeptide synthesis only took place at acidic pH values. Next, other individual components, 1) ammonia, in the form of NH_4Cl, 2) amino acid amide, and 3) amino acid nitrile, were added separately to the solutions to better understand the synthetic route to dipeptides in the cyanamide spark-discharge experiment and to evaluate the possible roles of these species in facilitating polymerization. These solutions were analyzed for dipeptides and DKPs after being subjected to heating at 50 °C for times of up to 3 weeks. Solutions that were not heated were frozen at 0 °C for use as a ($t = 0$) control.

The presence of ammonia resulted in negligible quantities of polymerization, so its role can be eliminated. However, it was observed that at a mildly basic pH value, cyanamide and dicyandiamide reacted readily in the presence of the amino

acid amide, and a factor of 2–4 times less in the presence of the amino acid nitrile, to generate dipeptides (see Figure S3 in the Supporting Information). These results indicate that the presence of the amino acid amide, or amino acid nitrile, is involved in the cyanamide-mediated amino acid polymerization reaction. It should be noted that under these conditions it was observed that dicyandiamide facilitated the formation of twice the concentration of dipeptides than did cyanamide, therefore dicyandiamide performed as the superior condensing reagent. This suggests that in the electric-discharge experiment, the dimerization of cyanamide, which is fastest at pH 9.6,[27] close to the pH value of the repeated cyanamide experiment, may have produced dicyandiamide within the discharge solution, where it then likely played a greater role in initiating amino acid polymerization, than cyanamide itself.

Scheme 1, which is based in part on other studies,[29] shows a possible mechanism for the cyanamide-mediated synthesis



Scheme 1. Scheme showing the dicyandiamide-mediated reactions involved in the polymerization of amino acids. The main dipeptide formation pathway is highlighted by the bold arrow, whereby the attack of the amino acid amide on the reactive intermediate (I) first yields the peptide amide, which is then hydrolyzed[30] to give the linear dipeptide (II).

of linear peptides at pH 9–10. Here, the carbodiimide form of cyanamide dimerizes to dicyandiamide, under mildly basic conditions, which can then be attacked by the nucleophilic carboxylate group of the amino acid to form the activated amino acid (I). At pH > 8, the amino group of glycinamide ($pK_a \approx 8$)[18] can attack the activated amino acid, the product of which can subsequently be hydrolyzed to ultimately give the linear dipeptide (II). Note that the pK_a value of the amino group of glycinamide is lower than that of the amino group of glycine, which is approximately 9.8.[18] As a consequence, the NH_2 group of glycinamide will be less protonated under such a regime, while the NH_2 group of glycine will be more protonated. Thus, glycinamide is a better nucleophile than free glycine in the pH regime of the spark-discharge experiment. However, it is worth noting that the unprotonated amino acid would also be a reactive species at pH values greater than the pK_a value of the amino group in glycine.

Hydrolysis of the amino acid amide to yield the amino acid is a potentially inhibitive pathway to dipeptide formation that should be considered. At pH = 9.75 and 55 °C, the half-life of glycinamide is about 3 days, while at pH = 7.95 and 75 °C, the half-life of glycinamide is about 7 days.[30] Extrapolating from these data, and considering that the cyanamide experiment was also mildly basic and that the reaction flask was no longer heated after the introduction of cyanamide (see the Supporting Information), glycinamide is expected to have had a sufficiently long lifetime to help facilitate the observed polymerization chemistry. Likewise, it is probable that the same is true in the case of the heating experiments that were performed that mimicked the spark-discharge solutions.

Also shown in Scheme 1 are several possible routes by which a second activated monomer (IA) can be formed as a by-product, which itself may undergo similar reactions as dicyandiamide to form the linear dipeptide. These possible additional dipeptide formation pathways may help explain why dicyandiamide induces more amino acid polymerization than does cyanamide.

The findings detailed here demonstrate the simultaneous synthesis of both simple and complex molecules under plausible prebiotic conditions. Miller's cyanamide experiment marks the first effort to study a prebiotic condensing reagent for its implications to the origin of life. Additionally, the results obtained here highlight the potential importance of condensing reagents in providing a mechanism to explain how simple organic compounds such as amino acids may have polymerized to form more complex molecules, such as dipeptides. The synthesis of dipeptides and DKPs by the cyanamide polymerization reaction may have additional implications, as some dipeptides and DKPs have been found to have catalytic properties that may have been important on the primordial Earth.[31]

Received: March 25, 2014
Published online: ■■ ■■, ■■■■

Keywords: amino acids · cyanamide · mass spectrometry · peptides · polymerization

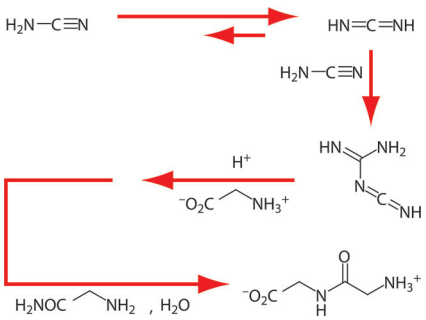
[1] S. L. Miller, *Science* **1953**, *117*, 528 – 529.
[2] H. C. Urey, *Proc. Natl. Acad. Sci. USA* **1952**, *38*, 351 – 363.
[3] J. F. Kasting, *Science* **1993**, *259*, 920 – 926.
[4] H. J. Cleaves, J. H. Chalmers, A. Lazcano, S. L. Miller, J. L. Bada, *Origins Life Evol. Biospheres* **2008**, *38*, 105 – 115.
[5] a) A. P. Johnson, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, A. Lazcano, J. L. Bada, *Science* **2008**, *322*, 404; b) E. T. Parker, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, M. Callahan, A. Aubrey, A. Lazcano, J. L. Bada, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5526 – 5531; c) F. Tian, O. B. Toon, A. A. Pavlov, H. De Sterck, *Science* **2005**, *308*, 1014 – 1017; d) J. C. G. Walker, P. Brimblecombe, *Precambrian Res.* **1985**, *28*, 205 – 222.
[6] a) K. Kvenvolden, J. Lawless, K. Pering, E. Peterson, J. Flores, C. Ponnamperuma, I. R. Kaplan, C. Moore, *Nature* **1970**, *228*, 923 – 926; b) O. Botta, J. L. Bada, *Surv. Geophys.* **2002**, *23*, 411 – 467; c) S. Pizzarello, G. W. Cooper, G. J. Flynn, *Meteorites and the Early Solar System II* (Eds.: D. S. Lauretta, H. Y. McSween, Jr.), The University of Arizona Press, Tucson, AZ, **2006**, pp. 625 – 651.
[7] a) J. L. Bada, *Chem. Soc. Rev.* **2013**, *42*, 2186 – 2196; b) A. S. Burton, J. C. Stern, J. E. Elsila, D. P. Glavin, J. P. Dworkin, *Chem. Soc. Rev.* **2012**, *41*, 5459 – 5472.
[8] a) T. Jakschitz, D. Fitz, B. M. Rode, *Cellular Origin, Life in Extreme Habitats and Astrobiology, Vol. 22* (Ed.: J. Seckbach), Springer Science and Business Media, Dordrecht, **2012**, pp. 469 – 489; b) G. Danger, R. Plasson, R. Pascal, *Chem. Soc. Rev.* **2012**, *41*, 5416 – 5429.
[9] M. A. Saito, D. M. Sigman, F. M. M. Morel, *Inorg. Chim. Acta* **2003**, *356*, 308 – 318.
[10] J. J. R. F. da Silva, R. J. P. Williams, *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*, Oxford University Press, Oxford, **1991**.
[11] L. Leman, L. Orgel, M. R. Ghadiri, *Science* **2004**, *306*, 283 – 286.
[12] A. J. Brearley, *Meteorites and the Early Solar System II* (Eds.: D. S. Lauretta, H. Y. McSween, Jr.), The University of Arizona Press, Tucson, Arizona, USA, **2006**, pp. 587 – 624.
[13] K. A. B. Dyl, A. K. Ziegler, E. D. Young, K. Wimmer, P. A. Bland, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 18306 – 18311.
[14] a) P. Schmitt-Kopplin, Z. Gabelica, R. D. Gougeon, A. Fekete, B. Kanawati, M. Harir, I. Gebefuegi, G. Eckel, N. Hertkorn, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2763 – 2768; b) G. D. Cody, C. M. O. D. Alexander, F. Tera, *Geochim. Cosmochim. Acta* **2002**, *66*, 1851 – 1865.
[15] A. Shimoyama, R. Ogasawara, *Origins Life Evol. Biospheres* **2002**, *32*, 165 – 179.
[16] a) A. Schimpl, R. M. Lemmon, M. Calvin, *Science* **1965**, *147*, 149 – 150; b) C. Ponnamperuma, E. Peterson, *Science* **1965**, *147*, 1572 – 1574.
[17] D. W. Nooner, E. Sherwood, M. A. More, J. Oró, *J. Mol. Evol.* **1977**, *10*, 211 – 220.
[18] R. N. Goldberg, N. Kishore, R. M. Lennen, *J. Phys. Chem. Ref. Data* **2002**, *31*, 231 – 370.
[19] G. Danger, A. Michaut, M. Bucchi, L. Boiteau, J. Canal, R. Plasson, R. Pascal, *Angew. Chem.* **2013**, *125*, 639 – 642; *Angew. Chem. Int. Ed.* **2013**, *52*, 611 – 614.
[20] M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239 – 242.
[21] M. W. Powner, S. L. Zheng, J. W. Szostak, *J. Am. Chem. Soc.* **2012**, *134*, 13889 – 13895.
[22] a) D. D. Christianson, J. S. Wall, R. J. Dimler, F. R. Senti, *Anal. Chem.* **1960**, *32*, 874 – 878: b) Miller's Laboratory Notebook 2, pages 110–112, Serial number 655, MSS642, Box 25, Mandeville Collections, Geisel Library, UCSD.
[23] B. R. Martin, *Biopolymers* **1998**, *45*, 351 – 353.
[24] A. Radzicka, R. Wolfenden, *J. Am. Chem. Soc.* **1996**, *118*, 6105 – 6109.
[25] J. Eveleigh, G. Winter, *Protein Sequence Determination* (Ed.: S. B. Needleman), Springer, Berlin, **1970**, pp. 91 – 123.
[26] S. L. Miller, *Biochim. Biophys. Acta* **1957**, *23*, 480 – 489.
[27] G. H. Buchanan, G. Barsky, *J. Am. Chem. Soc.* **1930**, *52*, 195 – 206.
[28] G. Steinman, R. M. Lemmon, M. Calvin, *Science* **1965**, *147*, 1574.
[29] J. Hulshof, C. Ponnamperuma, *Origins Life* **1976**, *7*, 197 – 224.
[30] L. Meriwether, F. H. Westheimer, *J. Am. Chem. Soc.* **1956**, *78*, 5119 – 5123.
[31] a) M. S. Iyer, K. M. Gigstad, N. D. Namdev, M. Lipton, *Amino Acids* **1996**, *11*, 259 – 268; b) M. Gorlero, R. Wieczorek, K. Adamala, A. Giorgi, M. E. Schinina, P. Stano, P. L. Luisi, *FEBS Lett.* **2009**, *583*, 153 – 156; c) A. L. Weber, S. Pizzarello, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12713 – 12717.

Communications

Prebiotic Chemistry

E. T. Parker, M. Zhou, A. S. Burton,
D. P. Glavin, J. P. Dworkin,
R. Krishnamurthy, F. M. Fernández,*
J. L. Bada* _____ ■■■■–■■■■

A Plausible Simultaneous Synthesis of Amino Acids and Simple Peptides on the Primordial Earth



Prebiotic polymerization: Archived samples from Stanley Miller's previously unreported 1958 cyanamide experiment were investigated to evaluate cyanamide-mediated amino acid polymerization under prebiotic conditions. Aqueous heating experiments indicate that in the presence of an amino acid amide, the dimerization of cyanamide under the mildly basic conditions of the spark-discharge experiment significantly enhances polymerization reactions.

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

A Plausible Simultaneous Synthesis of Amino Acids and Simple Peptides on the Primordial Earth**

Eric T. Parker, Manshui Zhou, Aaron S. Burton, Daniel P. Glavin, Jason P. Dworkin, Ramanarayanan Krishnamurthy, Facundo M. Fernández, and Jeffrey L. Bada**

anie_201403683_sm_miscellaneous_information.pdf

Supporting Information

Chemicals and Reagents

The chemical analyses discussed in this study were performed either at the NASA Goddard Space Flight Center (GSFC) or the Georgia Institute of Technology (GT). All glassware and sample handling tools used at GSFC and GT were thoroughly cleaned using polished water and then wrapped in aluminum foil and heated overnight, in air, at 500 °C. The GSFC analyses used Millipore Integral 10 water (18.2 MΩ cm, 3 ppb total organic carbon), while the GT analyses used Nanopure Diamond water (18.2 MΩ cm). The reagents used at GSFC were purchased from Sigma-Aldrich, or Fischer Scientific. Standard amino acid solutions (~10^{-4} M) were prepared by dissolving individual amino acid crystals (97-99 % purity) in water before being combined to allow for their measurement in a single chromatographic separation (GSFC analyses). Unless otherwise stated, all chemicals used at GT were purchased from Sigma-Aldrich, Chem-Impex International, Bachem Bioscience, or EMD Chemicals, with known purities ranging from 98 % to 99+ %. Stock dipeptide and diketopiperazine (DKP) standard solutions (~10^{-5} M) were freshly prepared by dissolving individual dipeptide and DKP crystals in Nanopure water prior to being mixed (GT analyses).

High temperature acid vapor hydrolysis (GSFC analyses) used 6 M doubly-distilled hydrochloric acid (ddHCl).[S1] Liquid phase acid hydrolysis (GT analyses) used analytical grade HCl. Cation exchange chromatography (GSFC analyses) was performed using 2 M sodium hydroxide (NaOH), 1.5 M ddHCl, and 2 M ammonium hydroxide (NH_4OH). The 2 M NH_4OH was prepared from NH_3 gas, while the 2 M NaOH was prepared by dissolving 32 g of NaOH in 0.4 L water.

Three reagents were prepared for derivatization of samples prior to analysis at GSFC: (1) 0.1 M sodium borate, (2) *o*-phthaldialdehyde/N-acetyl-L-cysteine (OPA/NAC), a fluorescent tag used to label primary amino groups and provide enantiomeric separation of chiral centers, and (3) double distilled 0.1 M hydrazine hydrate. Reagents were prepared as follows: (1) was prepared by dissolving solid sodium borate after heating at 500 °C for 3 hours in Millipore water; reagent (2) was prepared by dissolving 4 mg OPA in 300 µL Fisher Optima grade methanol, and mixing with 670 µL of 0.1 M sodium borate (pH 9) and 30 µL of 0.5 M NAC, and (3) was prepared as described elsewhere.[S2]

Three reagents were used for liquid chromatography – mass spectrometry (LC-MS) amino acid analysis at GSFC: (A) water, (B) methanol (Fisher Optima grade), and (C) 50 mM ammonium formate (adjusted to pH 8), with 8% methanol.[S2] Two reagents were used for dipeptide and DKP analysis of the original cyanamide samples at GT: (A) 10 mM ammonium acetate and (B) LC-MS grade acetonitrile. Reagent (A) was prepared by dissolving 770 mg of ammonium acetate crystal into 1 L of water. Two reagents were used for analysis of repeated cyanamide experimental samples and heating experiment samples: (A) 10 mM ammonium formate and (B) LC-MS grade acetonitrile. Reagent (A) was made by dissolving 631 mg of ammonium formate crystal into 1 L of water.

Sample Discovery and Experimental Setup

Miller's preserved samples were discovered in 2007 at the Scripps Institution of Oceanography, where Miller had stored samples from his original 1950s spark discharge experiments. Miller's original laboratory notebooks indicate that the cyanamide spark discharge experiment was conducted in 1958 using a version of his original spark discharge apparatus, described elsewhere,[S3] prior to collecting and saving the samples (S.L. Miller, 1958, Laboratory Notebook 2, page 110, Serial number 655, MSS642, Box 25, Mandeville Collections, Geisel Library). The initial experimental conditions included 300 mmHg CH_4, 250 mmHg NH_3, and 250 mL of water and the experiment was started by generating a spark discharge across two tungsten electrodes, using a Tesla coil, and gently boiling the water (S.L. Miller, 1958, Laboratory Notebook 2, page 110, Serial number 655, MSS642, Box 25, Mandeville Collections, Geisel Library). Information regarding the type of Tesla coil used by Miller for his electric discharge experiments, and its output, can be found elsewhere.[S4] The experiment ran for 17.5 hours before being stopped for 6 hours to allow for the introduction of 3.3 mg of cyanamide in 50 mL of water. The experiment was stopped by turning off the Tesla coil and removing the apparatus from the heat source. After the addition of cyanamide, the experiment was then resumed for another 12.5 hours. The experiment was restarted by turning the Tesla coil back on and without boiling the water. For the remainder of the experiment, the use of heat was discontinued, perhaps to minimize the risk of thermally decomposing cyanamide. The experiment was then stopped for 5 hours for a second addition of cyanamide (4.4 mg of cyanamide in 25 mL of water). Afterward, the experiment ran for another 29 hours before being stopped a third time to attempt a final introduction of 4.4 mg of cyanamide in 25 mL of water. However, Miller noted that he "could not pump ligand back into the apparatus" because of pressure build-up inside the reaction flask and as a result removed 1 L of gas from the apparatus to allow the experiment to continue without over-pressuring the apparatus, meanwhile saving 500 mL of the extracted gas (S.L. Miller, 1958, Laboratory Notebook 2, page 110, Serial number 655, MSS642, Box 25, Mandeville Collections, Geisel Library).

The experiment was then resumed after 4.5 hours, and ran for another 86 hours, at which time the experiment was ended by turning off the Tesla coil. An overview of this experimental design is provided in Table S1.

Table S1. Time sequence of Miller's cyanamide experimental design noting the duration of the experiment and the details of cyanamide introduction into the reaction flask.

Time (hours)	To Be Introduced Into Reaction Flask	Action
0	250 mL H_2O, 300 mmHg CH_4, 250 mmHg NH_3	Start spark, start heat
17.5	3.3 mg cyanamide in 50 mL H_2O	Stop spark, stop heat
23.5	-	Start spark, no heat
36	4.4 mg cyanamide in 25 mL H_2O	Stop spark
41	-	Start spark, no heat
70	4.4 mg cyanamide in 25 mL H_2O	Stop spark
74.5	-	Start spark, no heat
160.5	-	Stop spark

Miller first used ion exchange chromatography to isolate total amino acids, followed by subsequent separation of individual sample fractions using acid cation exchange chromatography.[S4, S5] Sample fractions were then evaporated until dry and archived until their discovery and analyses.

Original Miller Cyanamide Sample Analyses

Miller analyzed the products from his 1950s electric discharge experiments via paper chromatography with ninhydrin detection.[S6] Since then, advances in analytical chemistry have paved the way for researchers in the origins of life field to detect a wider range of chemistries, and lower abundances, of organic compounds generated by abiotic experiments. Current analytical techniques are much more sensitive and selective, and are capable of precisely quantifying a much larger range of amino acids and their enantiomeric abundances. The use of the OPA/NAC reagent coupled with sophisticated liquid chromatography and mass spectrometry techniques has proven to be a powerful approach for analyzing amino acids in the old Miller extracts.[S7, S8] Since OPA/NAC tags primary amino groups, secondary amino acids such as proline are not detected by these methods. Dipeptides in the samples can be readily characterized by a combination of liquid chromatography, ion mobility spectrometry, and high resolution tandem mass spectrometry techniques.

Amino acids were analyzed at GSFC using high performance liquid chromatography with fluorescence detection and triple quadrupole mass spectrometry (HPLC-FD/QqQ-MS). Dipeptides and DKPs were analyzed at GT using ultra performance liquid chromatography coupled to quadrupole-travelling wave ion mobility spectrometry-time of flight mass spectrometry (UPLC-Q-TWIMS-TOF-MS).

Aliquots of Miller's original cyanamide samples were prepared for amino acid analysis at GSFC, wherein half of each sample aliquot underwent acid hydrolysis, while the other half did not. Both the hydrolyzed and unhydrolyzed fractions were desalted.[S9] Amino acid data reported here were obtained from the analyses of the unhydrolyzed fractions. Prior to analysis, 10 µL of each sample solution were mixed with 10 µL of 0.1 M sodium borate, 5 µL of OPA/NAC, and derivatized for 15 minutes.[S10] Once complete, the derivatization reaction was then quenched with 75 µL of 0.1 M hydrazine hydrate and immediately analyzed by HPLC-FD/QqQ-MS.

Amino acid analysis was carried out using a Waters 2695 HPLC coupled to a Waters 2475 fluorescence detector and a Waters Quattro Micro API triple quadrupole mass spectrometer. Chromatography was performed using a 250 mm x 4.6 mm, 5 µm particle size Phenomenex Luna phenyl-hexyl column. The following mobile phases were used for chromatographic separation: (A) Millipore water, (B) methanol (Optima grade), and (C) 50 mM ammonium formate (pH 8) with 8% methanol. The fluorescence detector utilized an excitation wavelength of 340 nm and an emission wavelength of 450 nm. For chromatographic separation of target analytes, the following gradient was used: 0-5 min, 100 % C; 5-15 min, 0-83 % A, 0-12 % B, 100-5 % C; 15-22 min, 83-75 % A, 12-20 % B, 5 % C; 22-35 min, 75-35 % A, 20-60 % B, 5 % C; 35-37 min, 35-0 % A, 60-100 % B, 5-0 % C; 37-45 min, 100 % B; 45-46 min, 100-0 % B, 0-100 % C 46-55 min, 100 % C. The buffer flow rate used was 1 mL min^{-1} and the flow was split 90 % to the fluorescence detector and 10 % to the mass spectrometer. The QqQ-MS was operated in Q_1 scan mode, with a scan speed of 450 amu sec^{-1}. An electrospray ionization (ESI) ion source operating in positive ion mode was used to detect OPA/NAC amino acid derivatives in the 50-500 mass-to-charge (m/*z*) range. The ESI

settings used were the following: desolvation gas (N_2) temperature/flow rate: 350°C/650 L hr^{-1}, respectively; capillary voltage: 3.8 kV; cone voltage: 30 V. A description of the amino acids identified in Miller's original cyanamide samples is shown in Table S2.

A portion of Miller's original cyanamide spark discharge sample residues were dissolved in 500 µL of a 50:50 (v:v) mixture of methanol and water prior to direct dipeptide and DKP analysis by UPLC-Q-TWIMS-TOF-MS. Aliquots (50 µL) of the resuspended samples were dried under nitrogen in preparation for liquid-phase acid hydrolysis as described elsewhere.[S11] Afterward, any remaining HCl was evaporated with a dry nitrogen flow. The acid hydrolyzed sample residues were finally reconstituted in 50 µL of a 50:50 (v:v) mixture of methanol and water before analysis. Hydrolyzed sample extracts were used for comparison to samples that were analyzed directly, to determine if the peptides tentatively identified in the original samples were decomposed by hydrolysis, which would be indicative of the breaking of the peptide bond.

A Waters Acquity UPLC coupled to a Synapt G2 high definition mass spectrometer operating in positive ion mode (Waters, Milford, MA, USA) was used for dipeptide and DKP analyses. Chromatographic separations were performed using a 100 mm x 2.1 mm, 1.7 µm particle size Waters Acquity UPLC BEH amide HILIC column. The mobile phases used were (A) 10 mM ammonium acetate and (B) LC-MS grade acetonitrile. The eluent flow rate was 0.3 mL min^{-1}. A typical chromatographic gradient used was: 0-4 min, 100 % B; 4-5 min, 100-96 % B; 5-10 min, 96-45 % B; 10-13.5 min, 45-10 % B; 13.5-14.5 min, 10 % B; 14.5-15 min, 10-100 % B; 15-16 min, 100 % B. Column temperature was maintained at 60 °C and samples were held at 5 °C prior to analysis.

Peptides were identified following a two-step procedure. First, accurate mass survey UPLC-MS runs were screened against a custom-built MarkerLynx (v.4.1) database of accurate masses to identify peaks that matched dipeptide $[M+H]^+$ ions. A match tolerance of 10 mDa was used. The custom database was created by inputting dipeptide accurate masses calculated in Excel using SQLite database browser v. 1.3. Amino acids considered in potential dipeptides included asparagine, glutamic acid, serine, glycine, alanine, amino-*n*-butyric acid, α-aminoisobutyric acid, valine, glutamine, proline, threonine, arginine, histidine, lysine, leucine, isoleucine, phenylalanine, tryptophan, tyrosine, aspartic acid, and ornithine. Following this screening step, MS/MS experiments were conducted by scheduled quadrupole precursor ion selection followed by TWIMS ion mobility gas-phase separation, and transfer-cell collision-induced dissociation (CID) to fragment the peptides tentatively identified in the screening step. The experiments involved acquiring dipeptide product ion spectra in a retention time window of 0.5 minutes centered on the retention times observed in UPLC-TWIMS-MS runs carried out without precursor ion selection. An example of the resulting data collected that lead to dipeptide identification can be seen in Figure S1. Various energies were selected for CID experiments to ensure observing diagnostic fragment ions (e.g. y_1). These optimum energies were determined *a priori* using dipeptide standards. Scheduled MS/MS experiments helped optimize the duty cycle of the quadrupole by allowing it to focus on select precursor ions, as opposed to continuously cycling through a list of potential precursor targets. The TWIMS trap cell utilized a helium gas flow of 190 mL min^{-1}, while the main mobility cell used a nitrogen gas flow of 100 mL min^{-1}. The wave velocity was ramped from 700 to 2000 m s^{-1}, and the wave height from 22 to 30 V. Peptides were identified by elemental formula, and *de novo* sequencing using the PEAKS software (v. 5.3 Bioinformatics Solutions, ON, Canada). A tolerance of 10 mDa was used for matching fragment ions.

A dual ESI system was utilized for lock mass correction purposes. The primary ESI source was used and operated according to the following conditions: capillary voltage, 4 kV; sample cone voltage, 20 V; extraction cone voltage, 4 V; source temperature, 120°C; desolvation gas (N_2) temperature, 250°C; desolvation gas flow rate, 650 L hr^{-1}. A 0.5 mM sodium formate solution was used for daily calibration of the mass spectrometer. The calibration solution was prepared by mixing 1000 µL of 0.1 M sodium hydroxide, 900 µL of deionized water and 100 µL of formic acid. The calibration solution was then diluted to 20 mL using deionized water. Because minor drifts in the m/*z* scale may occur after initial calibration, a reference ESI source was used to provide an independent signal using leucine-enkephalin as the standard. The reference ESI source had the same settings as the main one, except that it used a capillary voltage of 2.8 kV. The TOF analyzer was operated in "V-optics mode", which utilized a reflectron to supply a full width at half maximum mass resolution of 21,000 based on the protonated ion of leucine-enkephalin at m/*z* 556.2771. The detector setting was 2575 V. The range acquired in all experiments was 50-1200 m/*z*. Glycine-containing dipeptides and DKPs detected in the original Miller samples are shown in Table S3 and Table S4, respectively. Non-glycine containing dipeptides and higher peptides that were tentatively identified, but not quantified in this study, are listed in Table S3.

Amino acid-dipeptide equilibrium was monitored and agreed with published values.[S12] However, equilibrium, alone, cannot explain the presence of dipeptides because the samples from Miller's H_2S experiment,[S8] which were simultaneously analyzed in an identical fashion, did not show the presence of dipeptides. The presence of dipeptides is also influenced by such factors as temperature and pH.[S13]

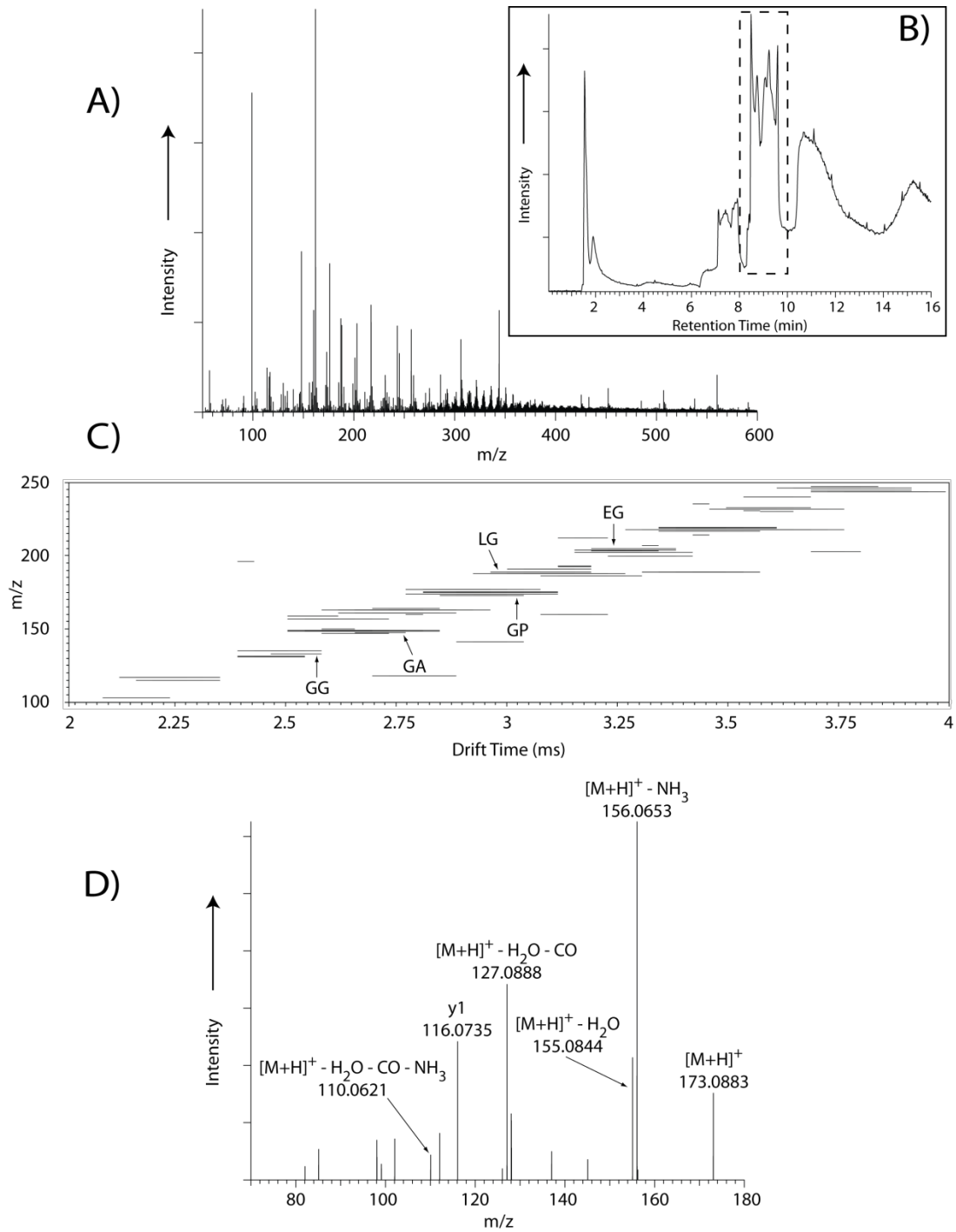


Figure S1. UPLC-Q-TWIMS-TOF-MS data demonstrating how dipeptides were identified in the archived cyanamide samples. Data from fraction 57-67 are shown here as an example. A) An averaged mass spectrum for the 8-10 minute region of the UPLC total ion chromatogram (B), exhibiting the vast chemical complexity of the samples studied here. C) TWIMS-TOF-MS m/*z* vs drift time plot for the 8-10 minute region of (B), where a suite of glycine-containing dipeptides are separated based on specific drift times. D) Following TWIMS separation, MS/MS spectra were obtained to confirm dipeptide detection by the presence of the $[M+H]^+$ ion, along with its corresponding fragment ions produced by collisional dissociation in the transfer cell; shown here is the MS/MS spectrum for glycylproline. Here, GG = glycylglycine, GA = glycylalanine, GP = glycylproline, LG = leucylglycine, EG = glutamylglycine.

Table S2. List of amino acids detected and quantitated in the archived cyanamide samples. Relative molar abundances of amino acids in the unhydrolyzed samples are normalized to alanine as opposed to glycine because the heavier alanine is farther away from the low molecular weight cut-off of the various mass spectrometry techniques used in this study and is thus detected with higher sensitivity. Given that a variety of mass spectrometers were employed to collect the data in this study, normalization to alanine provides for more reliable relative quantitation.

Amino Acid	STD LC RT (min)	Sample LC RT (min)	OPA/NAC Derivatized Mass $[M+H]^+$	STD Mass RT (min)	Sample Mass RT (min)	Molar Abundance (relative to Ala = 1)
D-Aspartic Acid	3.52	3.60	395	3.53	3.61	2.45×10^{-2}
L-Aspartic Acid	3.79	3.90	395	3.83	3.89	
L-Glutamic Acid	5.40	5.58	409	5.37	5.51	5.96×10^{-2}
D-Glutamic Acid	5.73	5.88	409	5.75	5.90	
D-Serine	11.84	11.81	367	11.80	11.82	2.49×10^{-2}
L-Serine	12.09	12.07	367	12.06	12.08	
Glycine	15.06	15.04	337	15.02	15.00	4.05
β-Alanine	15.62	15.62	351	15.58	15.58	0.88
D-Alanine	16.44	16.43	351	16.39	16.43	1
L-Alanine	16.72	16.71	351	16.69	16.69	
γ-ABA	16.44	16.43	365	16.41	16.36	7.16×10^{-3}
D-β-ABA	17.09	17.08	365	17.06	17.08	1.79×10^{-2}[a]
L-β-ABA	17.35	17.35	365	17.32	17.32	
D/L-α-ABA	17.83	17.83	365	17.78	17.78	1.57×10^{-2}
D-Isovaline	17.92	17.91	379	17.89	17.89	1.57×10^{-2}
L-Isovaline	18.00	18.00	379	17.96	17.96	
L-Valine	18.12	18.14	379	18.11	18.11	8.36×10^{-3}
D-Valine	18.39	18.39	379	18.35	18.37	
D-Isoleucine	18.94	18.94	393	18.91	18.91	8.24×10^{-4}
L-Isoleucine	19.42	19.40	393	19.39	19.35	
α-AIB[b]	17.35	17.35	365	17.32	17.32	-
D-Isoserine[b]	-	13.83	367	-	13.84	-
L-Isoserine[b]	-	14.32	367	-	14.30	-

[a] L-β-ABA and α-AIB interfered with one another during HPLC-FD/QqQ-MS analysis and were not quantitated. Due to the nearly racemic nature of major amino acids, the molar abundance of β-ABA is estimated here as twice the abundance of D-β-ABA. [b] Amino acids tentatively detected, but not quantified.

Table S3. Glycine-containing dipeptides detected in the original cyanamide samples.

Dipeptide	Theoretical $[M+H]^+/y_1$	Experimental $[M+H]^+/y_1$	$[M+H]^+/y_1$ Δm (mDa)	Molar Abundance (relative to Ala = 1)
Gly-Gly	133.0613/76.0399	133.0670/76.0468	5.7/6.9	6.96×10^{-4}
Gly-Ala	147.0770/90.0555	147.0772/90.0557	0.2/0.2	3.83×10^{-3}
Gly-Thr	177.0875/120.0661	177.0913/120.0676	3.8/1.5	5.26×10^{-4}
Gly-Pro	173.0926/116.0712	173.0883/116.0735	4.3/2.3	1.34×10^{-2}
Pro-Gly[a]	173.0926/76.0399	173.0920/-	0.6/-	2.65×10^{-3}
Gly-Val/Val-Gly	175.1083/(118.0868/76.0399)	175.1077/(118.0845/76.0353)	0.6/(2.3/4.6)	7.57×10^{-3}
Gly-Glu	205.0824/148.0610	205.0827/148.0596	0.3/1.4	2.56×10^{-3}
Glu-Gly	205.0824/76.0399	205.0829/76.0441	0.5/4.2	4.11×10^{-3}
Leu-Gly	189.1239/76.0399	189.1254/76.0416	1.5/1.7	6.78×10^{-4}
Ala-Pro[b]	187.1083	187.1084	0.1	-
Glu-Thr/Thr-Glu[b]	249.1087	249.1166	7.9	-
Glu-His[b]	285.1199	285.1179	2.0	-
Ser-Val[b]	205.1188	205.1187	0.1	-
Pro-Pro-Gly[b]	270.1454	270.1434	2.0	-
Asp-Asp-Gly[b]	306.0937	306.0980	4.3	-
Ser-Ser-Arg[b]	349.1836	349.1833	0.3	-

[a] The corresponding y_1 ion was not detected for Pro-Gly. [b] Dipeptides and tripeptides that were tentatively identified, but not quantitated.

Table S4. Glycine-containing DKPs detected and quantified in Miller's archived cyanamide experiment spark discharge samples. MS/MS analysis of fragmentation ions was not performed for glycine-containing DKPs.

DKP	Theoretical Mass $[M+H]^+$	Experimental Mass $[M+H]^+$	$[M+H]^+$ Δm (mDa)	Molar Abundance (relative to Ala = 1)
Cyclo(Gly-Gly)	115.0508	115.0602	9.4	0.24
Cyclo(Gly-Pro)	155.0821	155.0822	0.1	0.16
Cyclo(Leu-Gly)	171.1134	171.1133	0.1	0.10

Repeated Cyanamide Experiment Design

Duplication of Miller's cyanamide experiment was based on his original laboratory notebooks (S.L. Miller, 1958, Laboratory Notebook 2, page 110, Serial number 655, MSS642, Box 25, Mandeville Collections, Geisel Library). The duplicate experiment used heat and a spark in alternating one hour on/off cycles to preserve the lifetime of the spark generator and to minimize condensation of water vapor on the electrode tips, thereby mitigating the possibility of quenching the spark. Because of this cyclical sparking pattern, the time period over which the experiment was run was doubled to allow for an equivalent amount of spark reaction time as in the original cyanamide experiment. The initial experimental conditions were 250 mL of water, 300 mmHg CH_4 and 250 mmHg NH_3, all inserted into a 3 L reaction flask equipped with 2 tungsten electrodes. A spark was generated across the electrodes over a 14 day period. The times at which cyanamide was introduced into the reaction flask were also determined with the cyclical sparking pattern and Miller's original protocol in mind. General procedures used to prepare the repeated cyanamide experiment were based on other work.[S14]

An aqueous cyanamide solution was introduced into the 3 L reaction flask at three distinct times over the course of the experiment. To facilitate each cyanamide addition, the reaction was stopped by removing the apparatus from the heat and the spark for the same amount of time as Miller had originally done (Table S1). However, Miller's notebooks did not specify how the cyanamide introduction was performed; therefore an introduction apparatus was designed and implemented (Figure S2) to perform this task. Prior to each introduction, the aqueous cyanamide solution was loaded into a separatory funnel that was modified to be equipped with a cooling jacket, which was simultaneously connected to a vacuum system. A dry ice/acetone bath was placed in the cooling jacket surrounding the separatory funnel. The purpose of this cooling bath was to freeze the cyanamide solution and force dissolved gases out of the aqueous phase and into the above headspace of the separatory funnel. The headspace gas above the frozen solution was then evacuated and replaced with argon gas and evacuated again. The solution was then allowed to thaw at room temperature, prior to undergoing a second round of freezing, degassing, and thawing as described above. Once thoroughly degassed, the aqueous cyanamide solution was then introduced into the reaction apparatus using a sufficient pressure of Ar gas to push the aqueous solution into the reaction flask. Argon gas was chosen for the aforementioned uses, as opposed to alternative gases like N_2, because argon is not only chemically inert, but also would make it possible to avoid introducing any additional quantities of elements that were part of the initial experimental conditions.

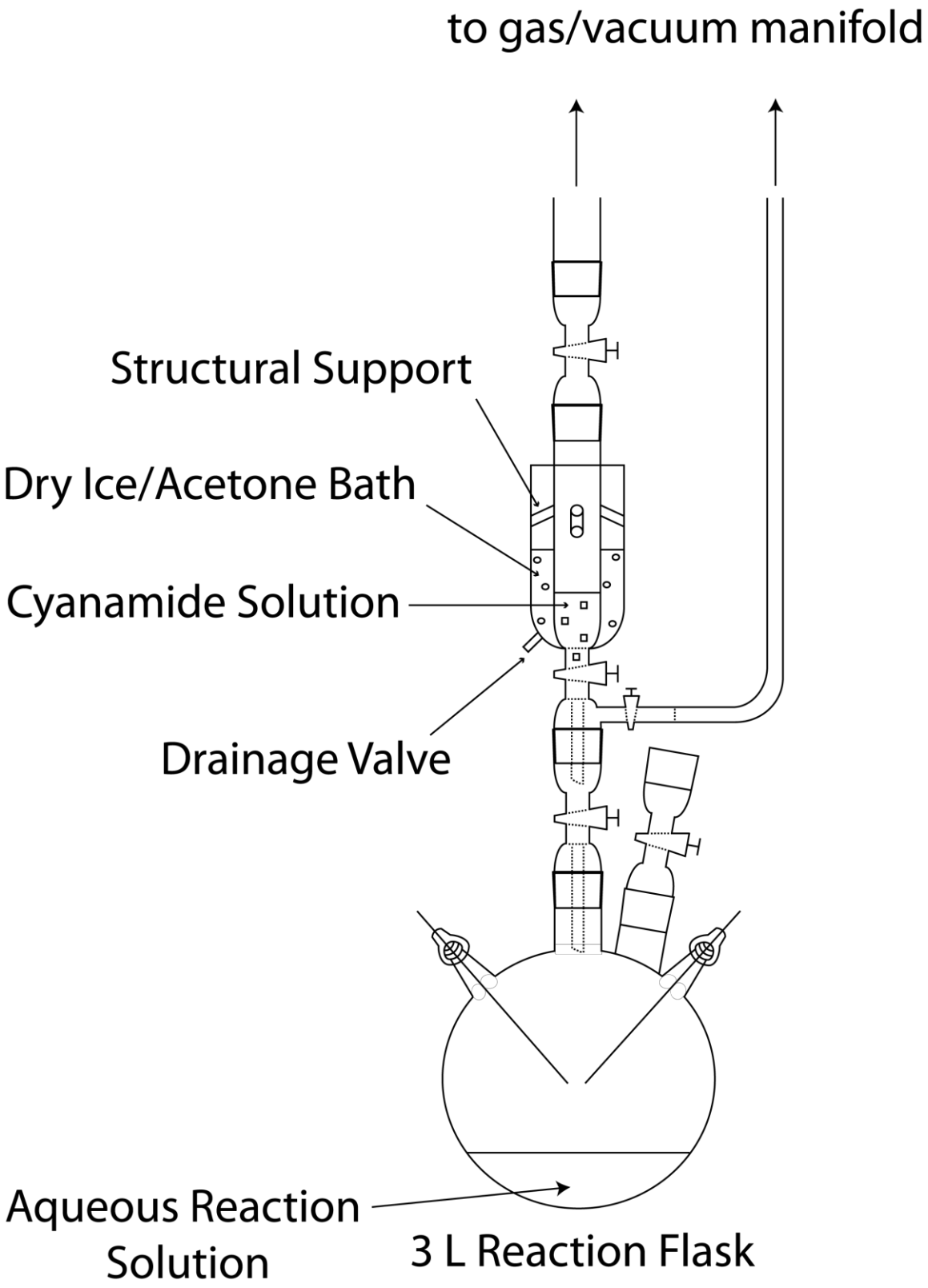


Figure S2. Schematic diagram of the setup used to conduct aqueous cyanamide introductions into the reaction apparatus without compromising the seal of, or introducing contaminants into, the flask while repeating Miller's cyanamide spark discharge experiment.

The first cyanamide addition came at t = 35 hours, where 3.3 mg of cyanamide dissolved in 50 mL H_2O was introduced into the 3 L reaction flask. The experiment was stopped for 6 hours for this cyanamide addition before restarting the experiment at t = 41 hours by resuming sparking and discontinuing heating the apparatus. The second aliquot of cyanamide was added at t = 66 hours by introducing 4.4 mg of cyanamide dissolved in 25 mL H_2O. The experiment was discontinued for 5 hours for this introduction before being resumed at t = 71 hours. The third and final cyanamide addition was at t = 129 hours where 4.4 mg of cyanamide dissolved in 25 mL H_2O was introduced. The experiment resumed at t = 133.5 hours and ran before finally being stopped at t = 305.5 hours. Liquid sample was carefully collected from the 3 L reaction flask, placed in sterilized receptacles, and stored in a freezer prior to analysis.

In addition to repeating Miller's cyanamide experiment, control and blank experiments were conducted for comparison to the repeated cyanamide experiment. The control and blank experiments were carried out in a fashion similar to the repeated cyanamide experiment. The control experiment differed from the repeated cyanamide experiment in that instead of introducing aqueous cyanamide solutions, only Nanopure water was introduced in volumes that matched those introduced during the repeated cyanamide experiment. This was done to determine if dipeptides could be detected without the presence of cyanamide. In the blank experiment, carbon was not introduced into the reaction flask. Instead, 300 mmHg N_2 was introduced, as opposed to 300 mmHg CH_4. Additionally, Nanopure water was introduced as was done for the control experiment, as opposed to aqueous cyanamide solutions.

Repeated Cyanamide Experiment Sample Analyses

Samples from the repeated cyanamide electric discharge experiment were analyzed by ultrahigh performance liquid chromatography (UHPLC)-QqQ-MS for confirmation of amino acid polymerization results. Aliquots (20 mL) of the repeated cyanamide experiment samples were lyophilized overnight. The resulting residues of this freeze-drying process were reconstituted in 2 mL of a 50:50 (v:v) mixture of LC-MS grade acetonitrile and 10 mM ammonium formate. The resulting supernatant was then sampled for direct analysis.

UHPLC-QqQ-MS experiments were performed with an Agilent 1290 Infinity UHPLC unit coupled to an Agilent 6430 QqQ-MS operated in multiple reaction monitoring mode. Chromatography was carried out with a 100 mm x 2.1 mm, 1.7 µm particle size Waters Acquity UPLC BEH amide HILIC column. The mobile phases used were A) 10 mM ammonium formate and B) acetonitrile (LC-MS grade). Linear dipeptides were eluted using isocratic runs of 30 % B. Isocratic elution of DKPs was achieved with a mobile phase composition of 85% B. The mobile phase flow rate was 0.3 mL min^{-1}, while the column was maintained at a temperature of 60 °C. The autosampler tray was kept at a temperature of 5 °C, the injection volume used was 5 µL, and the run time was 3 minutes. The QqQ-mass spectrometer was equipped with an electrospray ionization source that was operated in positive mode, with a capillary voltage of 4 kV, nebulizing gas (N_2) at a pressure of 15 psi and a desolvation gas temperature/flow rate of 300 °C/6 L min^{-1}, respectively .

Upon analyzing the repeat experiment samples, there was evidence to indicate that polymerization products had been formed. Among these biomolecules were some of the same peptides that were found in the original samples, including glycylglycine and 2,5-diketopiperazine. Some peptides not reported in the original samples were also detected in the repeated experiment samples. These included such organic molecules as alanylalanine and alanine anhydride. These compounds were identified based on retention times of parent masses corresponding to dipeptides and DKPs of interest.

Heating Experiment Design

Aqueous solutions were prepared for heating experiments designed to evaluate the polymerization of amino acids under a variety of pH regimes and a constant temperature. The temperature used for the experiments was 50°C and the pH ranges across several experiments were: a) pH 1-2, b) pH 6-7, c) pH 9-10, and d) pH 12-13. The pH 1-2 solutions were made of 0.1 M ddHCl. The pH 6-7 solutions were made by mixing 0.1 M potassium dihydrogenorthophosphate with 0.1 M potassium hydroxide (KOH) until a pH between 6 and 7 was achieved. The KOH used was of 91.3% purity in KOH (NOAH Technologies Corporation, San Antonio, Texas). The pH 9-10 solutions were made of 0.01 M sodium tetraborate decahydrate. The pH 12-13 solutions were made by mixing 0.1 M KOH with 0.1 M ddHCl until a pH between 12 and 13 was achieved. To each solution, glycine and either cyanamide, or dicyandiamide were added until the solution was 0.01 M in glycine and 0.05 M in cyanamide, or dicyandiamide. For each of the four tested pH ranges, control samples were created that contained 0.01 M glycine, but did not contain a condensing reagent.

To better simulate the chemical environment contained within the discharge apparatus, heating experiment solutions were prepared in addition to those described above. To the stock pH 9-10 solutions, ammonia was added in the form of NH_4Cl, until the solution became 0.01 M in glycine, 0.05 M in either cyanamide or dicyandiamide, and

0.01 M in NH_4Cl. Furthermore, a set of samples was prepared from the pH 9-10 stock solutions to contain 0.01 M glycinamide·HCl, or 0.01 M glycinonitrile·HCl. After the addition of glycinamide·HCl and glycinonitrile·HCl, these samples had a pH of ~8.5-9.

Aliquots of each solution were analyzed immediately upon being made, thereby representing samples that experienced 0 days of heating. Portions of the solutions that underwent heating were sampled periodically over the course of 3 weeks to evaluate how the polymerization chemistry varied temporally. Once collected, each heated sample was dried down and resuspended in a 50:50 (v:v) solution of acetonitrile and 10 mM ammonium formate prior to analysis.

Heating Experiment Sample Analyses

Heating experiment samples were analyzed for peptides using the retention times of parent masses associated with peptides of interest, as well as detection of corresponding fragment ions, as outlined for the repeated experiment samples. Analyzing the sample solutions that contained glycine and cyanamide or dicyandiamide showed that peptides formed almost exclusively in the pH 1-2 range. Samples that contained glycine, cyanamide or dicyandiamide, and NH_4Cl did not show evidence of polymerization under basic conditions, similar to the pH of the repeated cyanamide spark discharge experiment aqueous solution, which had a pH of ~10. However, heating experiment samples containing glycinamide or glycinonitrile did show evidence of polymerization at basic pH values (Figure S3). Over a 3 week period, upwards of ~7 μM of glycylglycine was synthesized in a basic solution that contained an amino acid, the corresponding amino acid amide or nitrile, and cyanamide or its dimer.[S15] A proposed mechanism for the cyanamide-mediated amino acid polymerization process is shown in Figure S4. From these experiments, it is evident that the amino acid amide is a more reactive intermediate to encourage peptide synthesis, than the amino acid and condensing reagent, alone. Furthermore, from these findings it can be concluded that dicyandiamide is a better condensing reagent than cyanamide.

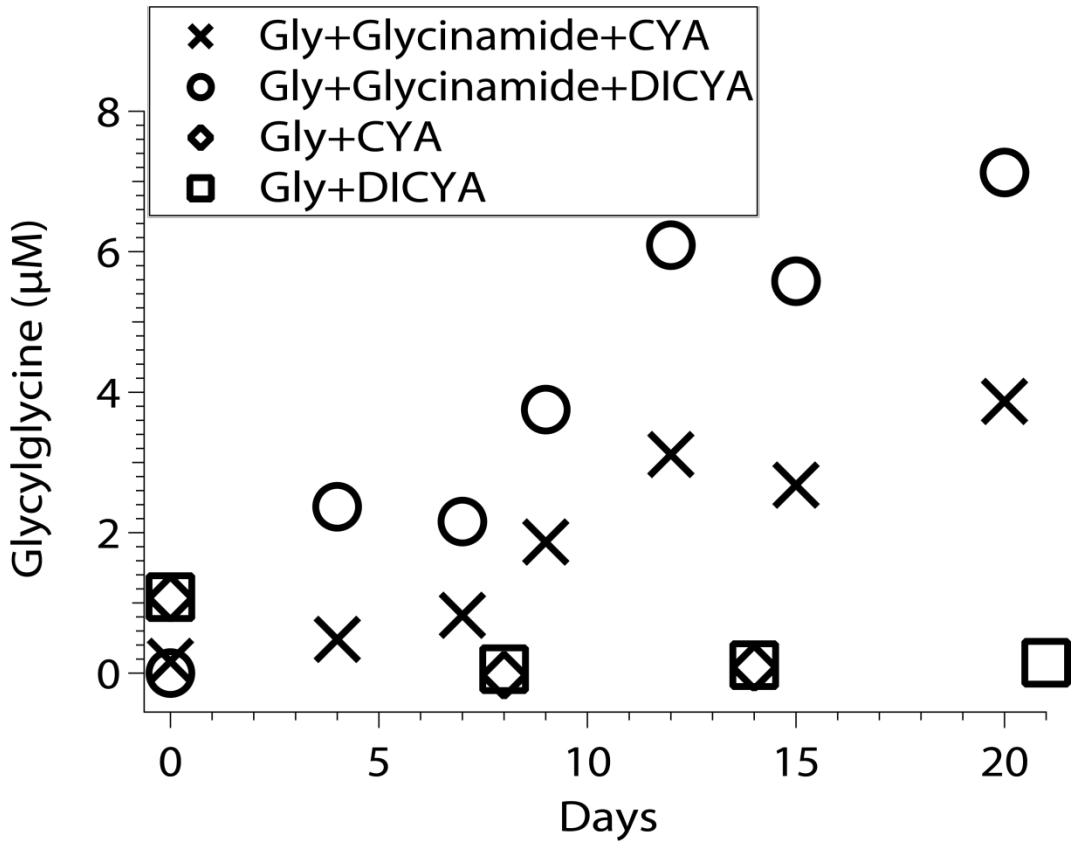


Figure S3. Change in glycylglycine concentration over time as observed from the analysis of heating experiment samples that contained glycine, and cyanamide (CYA) or dicyandiamide (DICYA), as well as glycine, glycinamide, and CYA or DICYA, in a basic solution. The pH values of the solutions shown here are as follows: Gly + CYA = 9.15, Gly + DICYA = 9.18, Gly + Glycinamide + CYA = 8.85, Gly + Glycinamide + DICYA = 8.42.

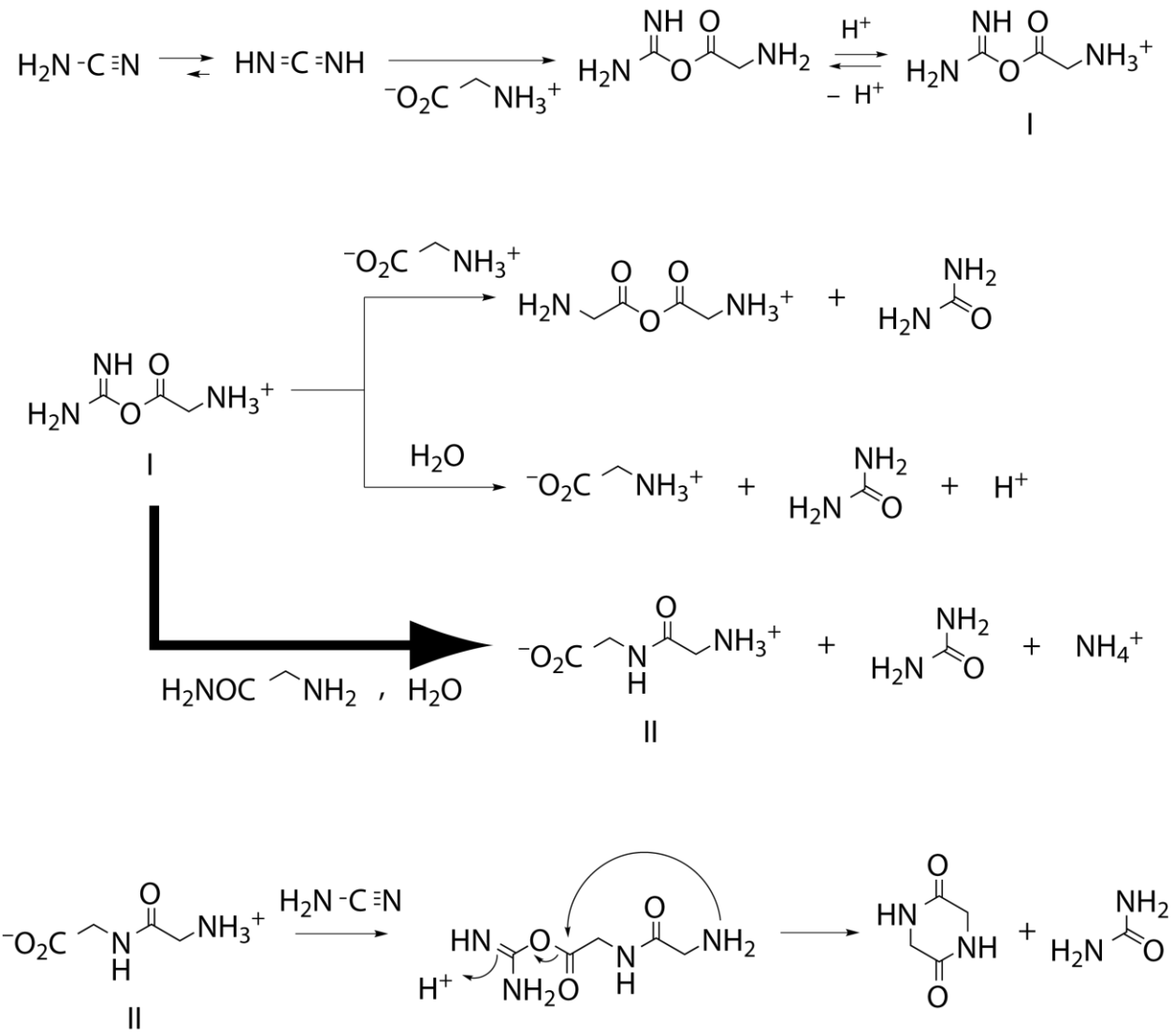


Figure S4. Scheme showing cyanamide-mediated amino acid polymerization reactions. Cyanamide reacts with an amino acid, glycine is shown as an example, to form an activated monomer (I). Once (I) is formed, it can undergo reaction with glycine to yield glycine anhydride, or with water to produce the original amino acid, among other by-products. The pK_a of the amino group in the amide, glycinamide is shown here as an example, is ~ 8 at pH values >~8,[S16] therefore the reaction of (I) with the amide is the favored reaction (bolded arrow) within a basic chemical environment to form the linear dipeptide (II). Product (II) can then react directly with cyanamide to undergo ring formation to yield the DKP. The scheme portrayed here is based in part on previously reported literature.[S15]

References

[S1] D. P. Glavin, A. D. Aubrey, M. P. Callahan, J. P. Dworkin, J. E. Elsila, E. T. Parker, J. L. Bada, P. Jenniskens, M. H. Shaddad, *Meteorit. Planet. Sci.* **2010**, *45*, 1695-1709.

[S2] D. P. Glavin, J. P. Dworkin, A. Aubrey, O. Botta, J. H. Doty, III, Z. Martins, J. L. Bada, *Meteorit. Planet. Sci.* **2006**, *41*, 889-902.

[S3] A. Lazcano, J. L. Bada, *Origins Life Evol. Biospheres* **2003**, *33*, 235-242.

[S4] S. L. Miller, *J. Am. Chem. Soc.* **1955**, *77*, 2351-2361.

[S5] D. D. Christianson, J. S. Wall, R. J. Dimler, F. R. Senti, *Anal. Chem.* **1960**, *32*, 874-878.

[S6] S. L. Miller, *Science* **1953**, *117*, 528-529.

[S7] A. P. Johnson, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, A. Lazcano, J. L. Bada, *Science* **2008**, *322*, 404.

[S8] E. T. Parker, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, M. Callahan, A. Aubrey, A. Lazcano, J. L. Bada, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 5526-5531.

[S9] A. S. Burton, J. C. Stern, J. E. Elsila, D. P. Glavin, J. P. Dworkin, *Chem. Soc. Rev.* **2012**, *41*, 5459-5472.

[S10] M. Zhao, J. L. Bada, *J. Chromatogr. A* **1995**, *690*, 55-63.

[S11] J. Eveleigh, G. Winter, *Protein Sequence Determination* (Ed.: Needleman, SB), Springer-Verlag, Berlin-Heidelberg-New York, **1970**, pp. 91-123.

[S12] B. R. Martin, *Biopolymers* **1998**, *45*, 351-353.

[S13] S. M. Gaines, J. L. Bada, *J. Org. Chem.* **1988**, *53*, 2757-2764.

[S14] E. T. Parker, H. J. Cleaves, A. S. Burton, D. P. Glavin, J. P. Dworkin, M. Zhou, J. L. Bada, F. M. Fernández, *J. Visualized Exp.* **2014**, *83*.

[S15] J. Hulshof, C. Ponnamperuma, *Origins Life* **1976**, *7*, 197-224.

[S16] R. N. Goldberg, N. Kishore, R. M. Lennen, *J. Phys. Chem. Ref. Data* **2002**, *31*, 231-370.