The Astrophysical Journal, 765:111 (9pp), 2013 March 10 doi:10.1088/0004-637X/765/2/111

ON THE FORMATION OF DIPEPTIDES IN INTERSTELLAR MODEL ICES

R. I. Kaiser[1], A. M. Stockton[2,3], Y. S. Kim[1], E. C. Jensen[2], and R. A. Mathies[2]
[1] Department of Chemistry, University of Hawaii at Manoa, Honolulu, HI 96822, USA
[2] Department of Chemistry, University of California, Berkeley, CA 94720, USA
[3] Jet Propulsion Laboratory, Pasadena, CA 01109, USA
Received 2012 October 10; accepted 2013 January 7; published 2013 February 25

ABSTRACT

The hypothesis of an exogenous origin and delivery of biologically important molecules to early Earth presents an alternative route to their terrestrial in situ formation. Dipeptides like Gly–Gly detected in the Murchison meteorite are considered as key molecules in prebiotic chemistry because biofunctional dipeptides present the vital link in the evolutionary transition from prebiotic amino acids to early proteins. However, the processes that could lead to the exogenous abiotic synthesis of dipeptides are unknown. Here, we report the identification of two proteinogenic dipeptides—Gly–Gly and Leu–Ala—formed via electron-irradiation of interstellar model ices followed by annealing the irradiated samples to 300 K. Our results indicate that the radiation-induced, non-enzymatic formation of proteinogenic dipeptides in interstellar ice analogs is facile. Once synthesized and incorporated into the "building material" of solar systems, biomolecules at least as complex as dipeptides could have been delivered to habitable planets such as early Earth by meteorites and comets, thus seeding the beginning of life as we know it.

Key words: astrobiology – astrochemistry – cosmic rays – ISM: molecules – methods: laboratory

Online-only material: color figures

1. INTRODUCTION

Molecular clouds are the nurseries of stars and planetary systems. Inside these clouds, grain particles at temperatures as low as 10 K accrete ice layers of water, methanol, carbon dioxide, carbon monoxide, ammonia, and methane (Gibb et al. 2004). These ices are processed chemically by intense ultraviolet (UV) radiation and by energetic galactic cosmic rays (GCRs) yielding complex organic molecules (Charnley et al. 2001). The densest parts of these clouds eventually undergo gravitational collapse to form primitive planetary material, which in turn supplies the basic ingredients for asteroids and cometary bodies. The significant deuterium enrichment in the organics found in carbonaceous chondrite meteorites demonstrates that these bodies contain a considerable fraction of pristine interstellar organic matter (Remusat et al. 2010). Therefore, organic material, such as amino acids, initially formed in the molecular cloud, can be incorporated into meteoritic parent bodies such as the Murchison meteorite (Shimoyama & Ogasawara 2002). Also, the amino acid glycine was detected recently in samples returned to Earth from comet 81P/Wild 2 by NASA's *Stardust* spacecraft (Elsila et al. 2009). Of crucial relevance to the prebiotic processes that preceded the onset of life on Earth is the formation of amino acid polymers, especially dipeptides. Once delivered to Earth by meteorites and comets (Oró 1961), dipeptides could have acted as catalysts in the formation of sugars and enzymes (Weber & Pizzarello 2006; Stano & Luisi 2009). Bio-functional polypeptides on an early Earth have been suggested to consist of up to four dipeptide blocks containing the amino acids aspartic acid (Asp), glycine (Gly), alanine (Ala), and valine (Val) (Ehrenfreund et al. 2001; van der Gulik et al. 2009). Laboratory experiments provide compelling evidence that amino acids can be formed abiotically via the Strecker synthesis and also by photolysis and charged particle processing of low-temperature interstellar ice analogs via ionizing radiation in the form of photons and energetic electrons (Bernstein et al. 2002; Muñoz Caro et al. 2002; Meinert et al. 2012; Holtom et al. 2005; Lafosse et al. 2006), but despite the vital role of dipeptides in astrobiology, the processes that could lead to the exogenous abiotic synthesis of a key class of astrobiologically relevant molecules in interstellar ices—dipeptides—are unknown.

Here, we explore the formation of dipeptides by subjecting interstellar model ices to ionizing radiation in the form of energetic electrons that model the impact of GCRs at 10 K. These ices contain hydrocarbons, carbon dioxide, and ammonia to address whether dipeptides can be synthesized abiotically from simple ice mixtures by interaction with ionizing radiation. The composition of our model ice is plausible based on the facts that (1) glycine can be formed during electron exposure of low temperature ices of methylamine (CH_3NH_2) with carbon dioxide (CO_2) and of acetic acid (CH_3COOH) with ammonia (NH_3), (2) carboxylic acids and amines can be formed in ices of hydrocarbons with carbon dioxide and ammonia, respectively, (3) hydrocarbons can be synthesized from methane (CH_4), and (4) methane, ammonia, and carbon dioxide are constituents of interstellar ices (Holtom et al. 2005; Lafosse et al. 2006; Kaiser & Roessler 1998; Kim & Kaiser 2010, 2011; Bennett & Kaiser 2007).

2. EXPERIMENTAL

The simulation experiments were conducted in a contamination-free ultra-high vacuum (4×10^{-11} torr) chamber (Bennett et al. 2004). Four sets of ice mixtures were prepared on a 10 K silver target via deposition of the premixed gases. These mixtures are (1) carbon dioxide (CO_2), ammonia (NH_3), and hydrocarbons (C_nH_{2n+2}; $n = 1$–6) methane, ethane, propane, *n*-butane, *n*-pentane, and *n*-hexane [18:8:1:1:1:1:1:1; complex hydrocarbon mixture], (2) carbon dioxide (CO_2), ammonia (NH_3), and the hydrocarbons methane, ethane, and propane [5:3:2:2:1; simple hydrocarbon mixture], (3) carbon dioxide (CO_2), ammonia (NH_3), and perdeuterated hydrocarbons methane, ethane, and propane [5:3:2:2:1; partially deuterated mixture], and (4) carbon dioxide (CO_2), D-ammonia (ND_3), and perdeuterated hydrocarbons methane, ethane, and propane [5:3:2:2:1; fully deuterated mixture]. Complex (1) and simple

(2) hydrocarbon mixtures were chosen to link the structures of the amino acids formed to their hydrocarbon precursors; partially (3) and fully (4) deuterated mixtures were selected to address the assignment of the infrared absorptions based on the isotopic shifts. The 10 K target temperature represents typical temperatures of ice-coated grains in cold molecular clouds; the pressure conditions of 4×10^{-11} torr guarantee that over the time scale of each experiment of 15 hr, less than one monolayer of residual gases condensed on the icy target. The 200 ± 20 nm thick ices were irradiated with energetic electrons (5 keV) for 60 minutes with $(1.5 \pm 0.2) \times 10^{12}$ electrons $cm^{-2}s^{-1}$. This experiment mimics the exposure of the ices for about 8×10^6 years (Ehrenfreund et al. 2001; van der Gulik et al. 2009), i.e., typical lifetimes of cold molecular clouds. After irradiation, the ices were heated at a rate of 0.5 K $minute^{-1}$ to 300 K. A Fourier transform infrared spectrometer (Nicolet 6700 FTIR) was used to follow the evolution of the ices. Note that no laboratory experiment can exactly simulate the interaction of the energetic GCRs with ices directly since no experimental device can generate a broad range of kinetic energies of protons and helium nuclei—the main constituents of the GCR—from the MeV to the PeV range. However, the physical effects of GCRs interacting with ices are understood and can be modeled: GCRs lose energy predominantly via ionization of the target molecules; the secondary electrons generated can induce further ionization, thus creating electron cascades (Drouin et al. 2001). By convolving over the energies of the GCR particles, it is feasible to derive a kinetic energy distribution of the electrons generated that are typically in ranges of a few eV up to 10 keV (Kaiser & Roessler 1997). Therefore, we simulate the GCRs by irradiating the samples with electrons, here at 5 keV. Their linear energy transfer (LET) is similar to LETs of 10–20 MeV GCR protons penetrating ices. The energy transfer from the 5 keV electrons to the ices was simulated utilizing the CASINO code to be about 1.0 ± 0.1 keV per electron (Drouin et al. 2001). This value relates to an average LET of 5.0 ± 0.5 keV μm^{-1}, i.e., comparable to a few keV μm^{-1} which 10–20 MeV GCRs deposit into interstellar ices (Kaiser & Roessler 1997). Each molecule in the ice mixtures is exposed to an average of up to 14 ± 2 eV. Each second in the laboratory simulates the exposure of $7 \pm 2 \times 10^{10}$ s in space. Therefore, the complete experiment mimics the exposure of the ices of about 8×10^6 years (Ehrenfreund et al. 2001; van der Gulik et al. 2009), i.e., typical lifetimes of cold molecular clouds. Thirdly, we cannot quantify to what extent the amino acids and dipeptides were formed in our experiments during the irradiation at 10 K or in the course of the annealing period via radical recombination. However, based on the infrared spectroscopy, the peptide bond is present in the 10 K irradiated ices as well as in the 300 K residues.

Due to the low sample mass and low conversion of interstellar ice analogs into amino acids and dipeptides, the highly sensitive technique of microchip capillary electrophoresis (μCE) with laser-induced fluorescence was used, providing much better sensitivity than contemporary GCMS approaches (Navarro-González et al. 2010). Here, the samples were analyzed for amino acids and dipeptides using the Mars Organic Analyzer (MOA), a μCE instrument designed for ultrasensitive detection and analysis of small organic biomarkers in the solar systems. The analytes were separated according to their electrophoretic mobility and identified based on spiking experiments using the authentic amino acids and dipeptides (Skelley et al. 2005). After the irradiated ices were heated to 300 K, the remaining organic residues were dissolved in a 1:1 mixture of ultra-clean buffer (100 μM borate) and 2 mM Pacific Blue succinimidyl ester without hydrolyzing potential dipeptides (Stockton et al. 2009). The samples were recovered from the silver targets using a combination of mechanical and chemical release. First, 20 μL of a reaction solution containing 2 mM Pacific Blue succinimidyl ester was prepared in DMF. The reaction solution was pipetted onto the target in 5 μL aliquots, allowed to react while undergoing agitation for five minutes and then transferred to a clean microtube. The surface film was mechanically roughened using a freshly sterilized razor blade between reaction solution treatments. The final volume recovered (17–19 μL) was measured, transferred to a fresh tube, and an equal volume of aqueous 100 μM recrystallized borate buffer, pH 9.5, was added. A sterile technique was used at all times in a laminar flow hood rated for work with RNA. After overnight reaction, the samples were diluted 1:1 with ultra-clean, 30 mM recrystallized borate buffer. Samples were then run with no further dilution on the Mars Organic Analyzer μCE instrument (Stockton et al. 2009; Skelley et al. 2005). Authentic amines, amino acids, and dipeptides were fluorescently labeled and spiked into the samples individually to confirm electrophoretic migration times of all assigned peaks.

3. RESULTS AND DISCUSSION

The infrared spectra of the organic residues that remain upon warming the irradiated sample to 300 K (Figure 1) were used to characterize the formation of newly emerging functional groups. First, we observe intense absorptions at 1667 cm^{-1} in the complex (Figure 1(A)) and 1687 cm^{-1} in the simple (Figure 1(B)) hydrocarbon-based mixtures that are assigned as the carbonyl stretching of the amide group (Table 1); in proteins, the amide I band is the most intense absorption band. Likewise, we observe the absorptions at 1670 cm^{-1} in the partially (Figure 1(C)) and 1657 cm^{-1} in the fully (Figure 1(D)) deuterated samples. Second, the following absorptions also stand out at 1590 cm^{-1} (Figure 1(A)) and 1611 cm^{-1} (Figure 1(B)) in hydrocarbon-based mixtures. These absorptions could correspond to the weak amide II band of a motion combining both the N–H bending and the C–N stretching vibrations or to the asymmetric stretch of deprotonated carboxylic acids (RCO_2^-); in the partially (3) and fully deuterated (4) samples, this band is observed at 1606 cm^{-1} and does not shift at all between samples. Therefore, this observation is inconsistent with the sole amide II band assignment. The amide III band is typically much weaker than the amide I and II absorptions and could be obscured by the broadbands at about 1380 cm^{-1} (Figure 1). The broad absorptions at 3236 cm^{-1} (Figure 1(A)) and 3200 cm^{-1} (Figure 1(B)) can be attributed to the N–H-stretching mode. In the partial mix (Figure 1(C)) a broadband remains at $\sim$3100 cm^{-1}, while in the fully deuterated mix (Figure 1(D)) the N–D-stretch shifts down to $\sim$2400 cm^{-1}. Prominent absorptions of aliphatic $–CH_3$ and $–CH_2–$ groups are visible at 2964/2965 cm^{-1}, 2929/2929 cm^{-1}, and 2870/2869 cm^{-1} (Figures 1(A) and (B)). These bands shift upon deuteration to 2219/2221 cm^{-1} (Figures 1(C) and (D)). The deuteration also diminishes the intensities of the $–CD_2–$ group. The deconvoluted bands at 2952/2945 cm^{-1} and 2769/2780 cm^{-1} (Figures 1(A) and (B)) can also be linked to the NH_4^+ and $–NH_3^+$ groups (Table 1); deuteration leads to the absorptions of ND_4^+ and $–ND_3^+$ at 2190 and 2089 cm^{-1} (Table 1). The non-deuterated samples have additional broad features in the 1460 to 900 cm^{-1} range, which could be linked to deformation and scissor modes of aliphatic $–CH_3$ and

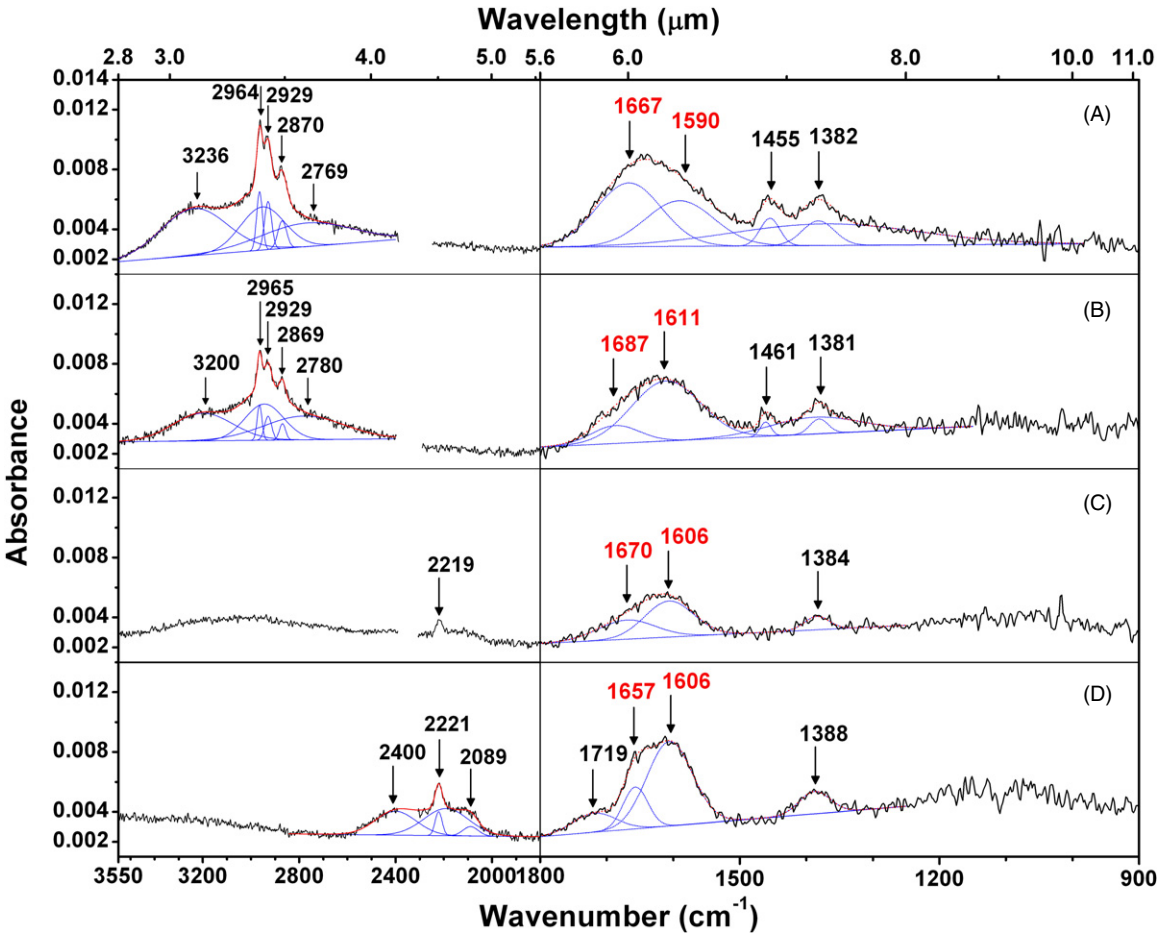

Figure 1. Mid-infrared spectra of the organic residues formed upon warming the irradiated ices mixtures to 300 K (black) and the sum of the band deconvolution (red). From top to bottom: (A) complex hydrocarbon-bearing mixture, (B) simple hydrocarbon-bearing mixture, (C) simple partially deuterated sample, and (D) simple fully deuterated sample. Absorptions originating from amide modes are highlighted. Minor contributions from atmospheric carbon dioxide at ~2400 cm^{-1} were traced out. The band assignments are compiled in Table 1.

(A color version of this figure is available in the online journal.)

–CH_2– groups. Finally, the band in the 1388–1371 cm^{-1} region does not shift upon deuteration; it can be attributed to the symmetric stretching mode of RCO_2^-. The weak 1719 cm^{-1} absorption seen in the fully deuterated sample (Figure 1(D)) appears at the correct position for carbonyl stretch (C=O) of the –COOD group. The corresponding absorptions of –COOH group in the other samples are expected to be present, but obscured by the high frequency side of the strong amide I bands.

Therefore, the infrared data reveal the formation of nonvolatile molecules containing functional groups such as amides (–CONH–/–CONR–), aliphatic groups (–CH_2–, –CH_3), (deprotonated) carboxylic acids (RCO_2^-), and protonated amines/ammonia (–NH_3^+, NH_4^+). However, these data alone cannot identify individual dipeptides or amino acids. We thus turned to the complementary, highly sensitive technique of μCE with laser-induced fluorescence (LIF) detection to identify individual amino acids and dipeptides in the residues. The primary amine-containing products in the room-temperature residues were labeled with the fluorescent dye Pacific Blue succinimidyl ester and electrophoretically separated utilizing μCE to produce the electropherograms shown in Figure 2. Nine amino acids were readily identified in all of the sample electropherogram traces based on their mobility and confirmed by direct spiking of the samples with authentic material (*Gly*, *Ala*, β-Ala, α-, β-, and γ-amino butyric acid, *Val*, nVal, and *Leu*); those indicated in italics are proteinogenic amino acids. The fact that these amino acids were produced by ice irradiation was ascertained by their significant and consistent appearance in the electropherogram traces of the irradiated samples compared to full procedural controls. A quantitative analysis indicates that the proteinogenic amino acids Gly and Ala account for up to 80% of the amino acids by mass in the residues with Gly dominating by a factor of about three compared to Ala; about 0.2% of the precursor molecules after control correction were converted into amino acids (see the Appendix). The nonvolatile material contained six amino acids with branched and non-branched hydrocarbon building blocks having one (Gly), two (Ala, β-Ala), and three (α,β, and γ-amino butyric acid) carbon atoms, which could be linked to the structures of the methane, ethane, and propane hydrocarbon precursors. Val and Leu contain branched hydrocarbon moieties; the proteinogenic amino acids Val and Leu can be derived from Gly and Ala by replacing a hydrogen atom at the C1 and C2 atoms, respectively, by the 2-propyl radical (H_3CCHCH_3) formed upon radiolysis.

Most important, in addition to the readily identified amino acid peaks, a significant number of new Pacific Blue labeled molecules appeared in the 120–140 s time range characteristic of dipeptides compared to non-irradiated samples (Figure 2). Peaks that co-migrate with Leu–Leu and Gly–Leu spikes are similarly strong in the controls and in all samples and are thus

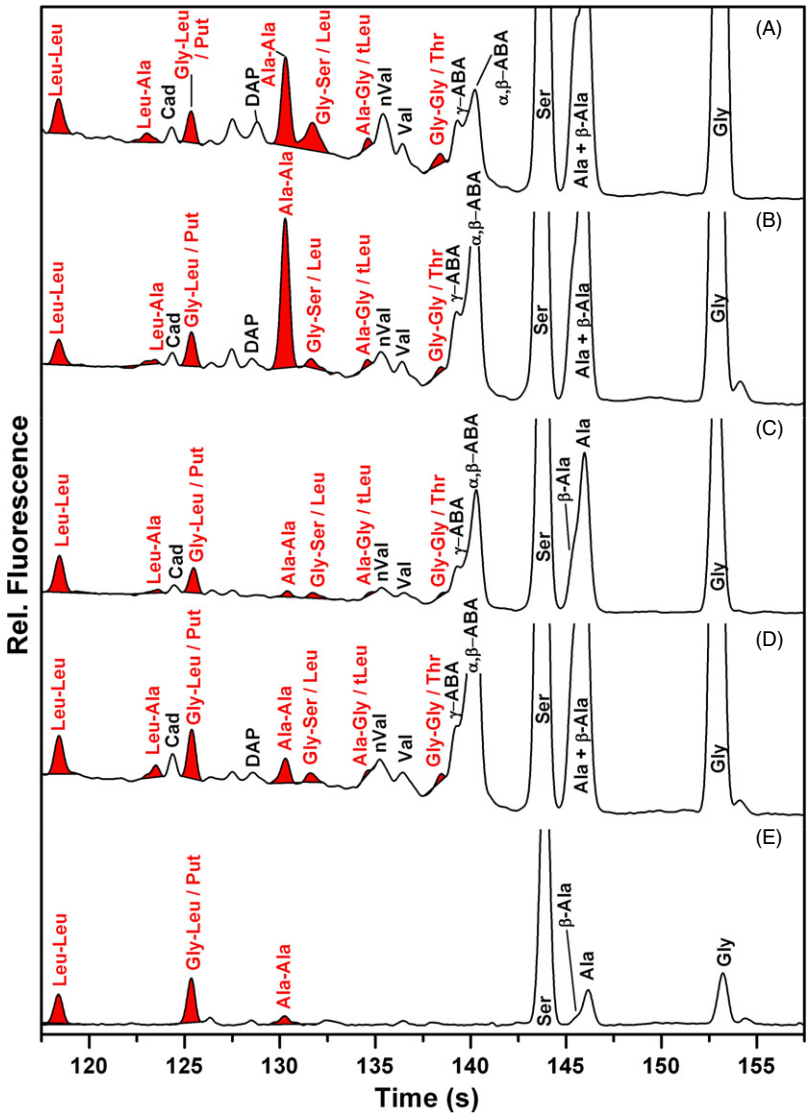

Figure 2. Microchip capillary zone electrophoretic (CZE) analysis of Pacific Blue labeled amino acids and dipeptides extracted from the room temperature organic residue produced by irradiation of interstellar model ices. From top to bottom: (A) complex hydrocarbon-bearing mixture, (B) simple hydrocarbon-bearing mixture, (C) simple partially deuterated mixture, (D) simple fully deuterated mixture, and (E) full procedural control. DAP: diaminopropane; Cad: cadaverine; Put: putresceine. Peaks that migrate coincident with the indicated dipeptides are shaded red. All traces are presented on the same relative sensitivity scale and offset for display.

(A color version of this figure is available in the online journal.)

uninformative. Peaks that co-migrate with Gly–Gly, Ala–Gly, Ala–Ala, and Leu–Ala spikes are consistently stronger in all irradiated samples (see Figure 2 and the Appendix). Of these, the peak migrating as the dipeptide Leu–Ala could be unambiguously assigned. The productions of Gly–Gly and Ala–Gly are less quantitative because these dipeptides co-migrate with Thr and tLeu, respectively (Figure 2). As expected from the identification of their formal amino acid precursors, dipeptides that are rich in amino acids Gly and Ala are synthesized preferentially (Figure 3). Quantitatively, the yields of dipeptides are about a factor of ten lower than the yields of the amino acid monomers. These observations suggest that once amino acids are formed in our experiment, ionizing radiation further radiolyzes the amine and carboxylic acid functional groups to trigger bond ruptures and to form reactive radicals as depicted in Figure 4.

What are the underlying reaction pathways to form these dipeptides? Previous mechanistic studies on the formation of the simplest amino acid glycine suggested two formation routes starting with methylamine (CH_3NH_2) and carbon dioxide (CO_2) as well as acetic acid (CH_3COOH) and ammonia (NH_3), respectively, upon interaction of low temperature ices with ionizing radiation (Holtom et al. 2005; Lafosse et al. 2006; Figure 5). Here, ionizing radiation interacts with methylamine leading to a hydrogen atom loss and formation of the CH_2NH_2 radical (R1); the hydrogen atom has excess kinetic energy and can overcome the barrier to addition to carbon dioxide yielding the HOCO radical (R2). The latter can recombine within the matrix cage with the CH_2NH_2 radical forming glycine. Alternatively, acetic acid interacts with ionizing radiation and releases a hydrogen atom from the methyl group leading to the CH_2COOH (R3). Upon interaction with ionizing radiation, ammonia can also lose atomic hydrogen yielding the NH_2 radical (R4). Both CH_2COOH and NH_2 recombine forming glycine. In summary, both routes require an amine plus carbon dioxide as well as a carboxylic acid plus ammonia, respectively, to form glycine. As indicated in Figure 5, this reaction scheme can be expanded from the simplest amine and carboxylic acid to more

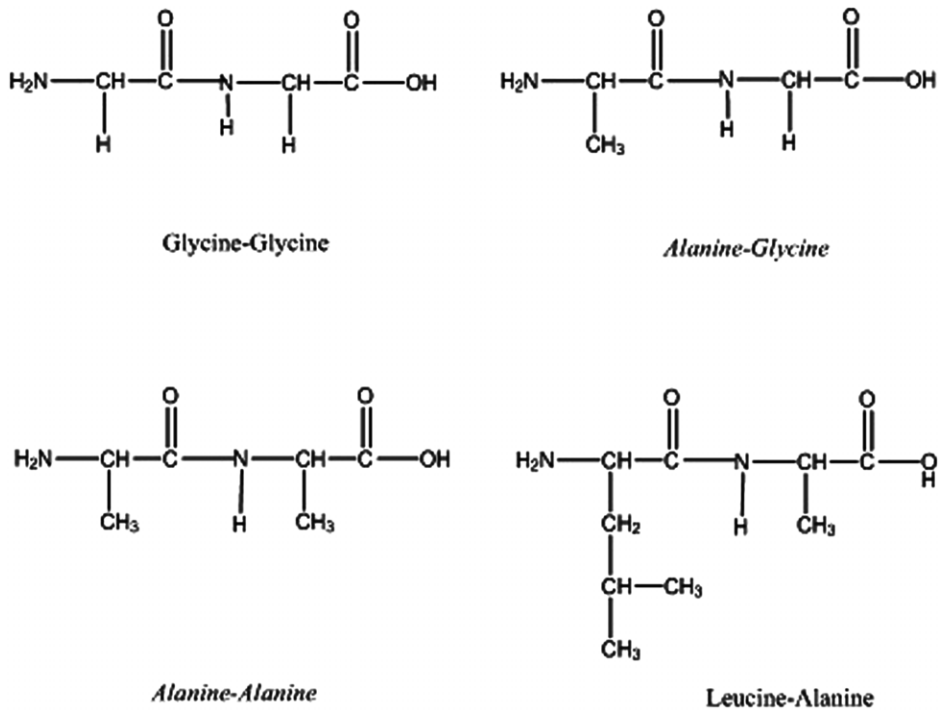

Figure 3. Dipeptides characterized via microchip capillary electrophoresis.

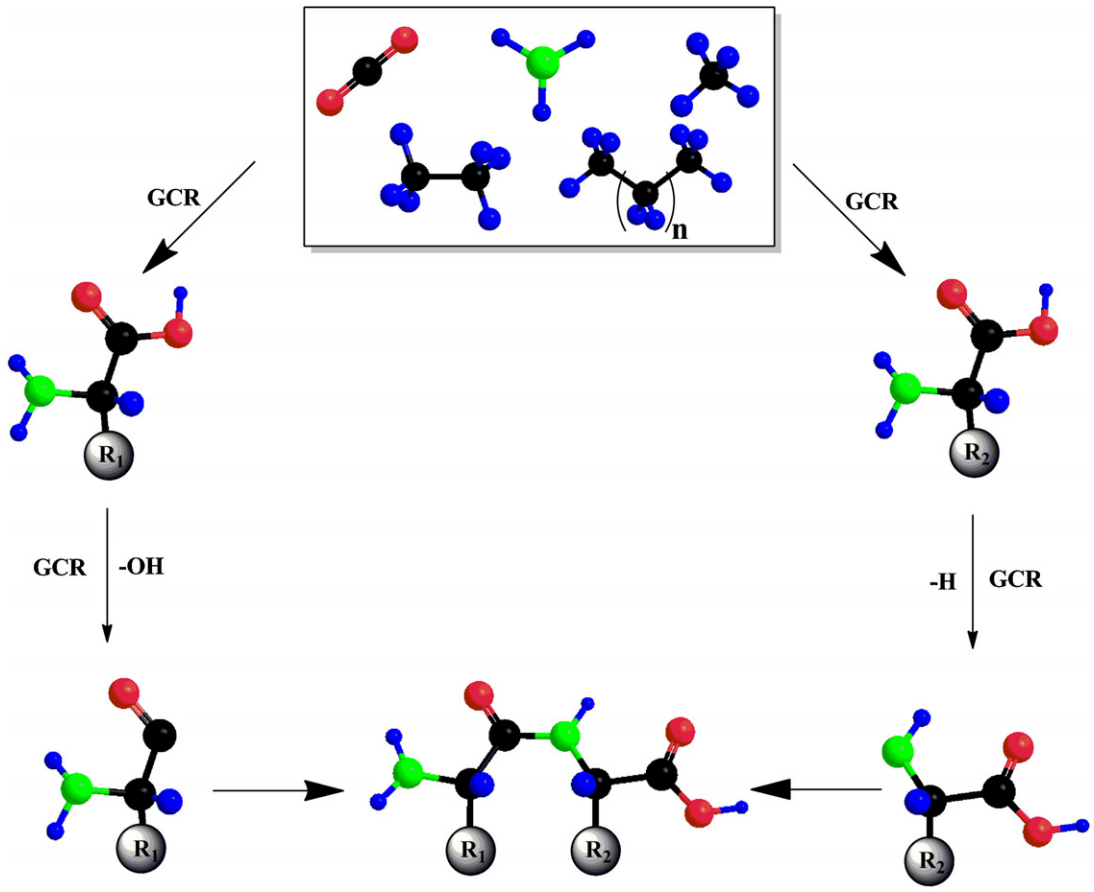

Figure 4. Schematic representation of chemical reactions leading to amino acids and dipeptides when interstellar ice models consisting of carbon dioxide, ammonia, and hydrocarbons are exposed to energetic electrons. GCR defines the galactic cosmic ray, which can lead to the loss of a hydroxyl radical or hydrogen atom from the amino acids. Radical–radical recombination either at 10 K or during the warm-up phase can yield the dipeptide. R1 and R2 define the side chains in the generic amino acids.

(A color version of this figure is available in the online journal.)

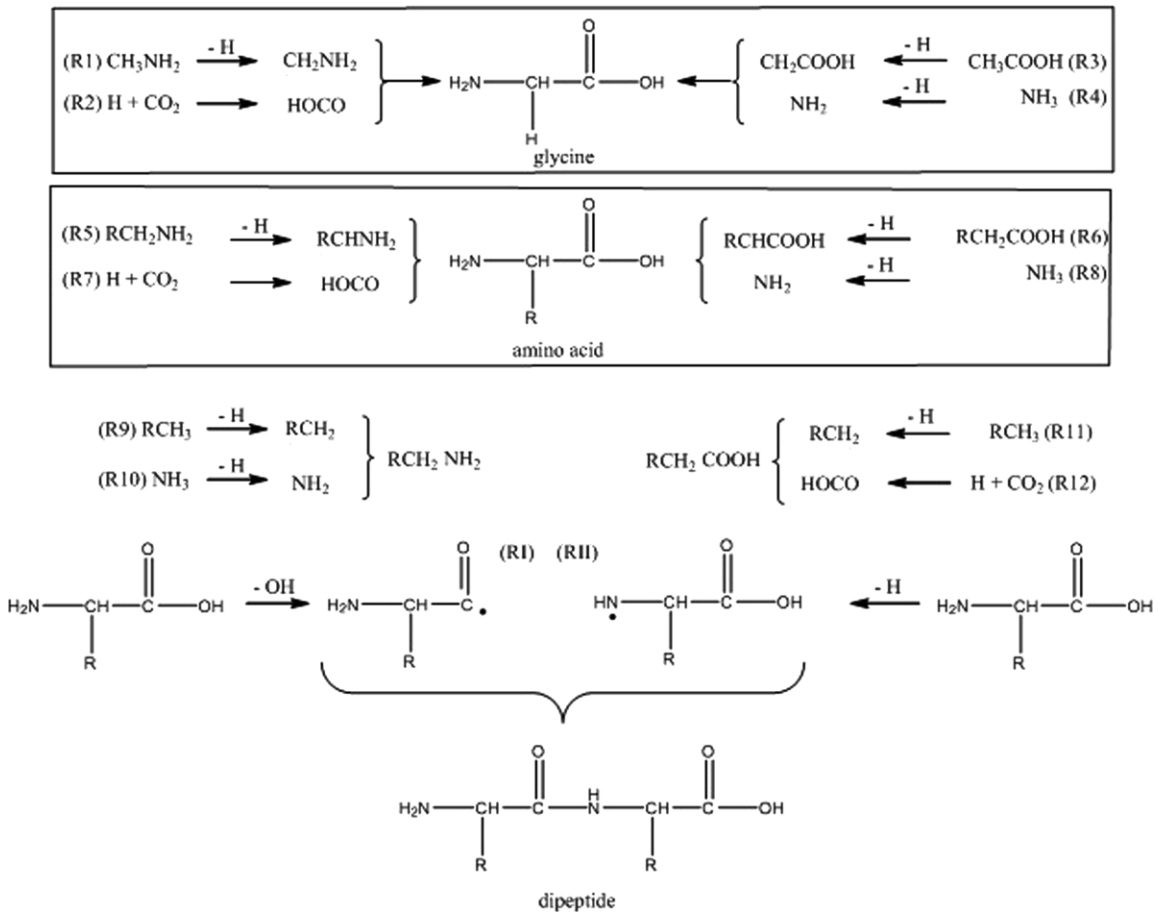

Figure 5. Proposed formation routes of amino acids and dipeptides in the radiolysis of complex ice mixtures upon interaction with ionizing radiation.

complex systems, which are schematically represented as RCH_2NH_2 and RCH_2COOH, respectively, with "R" representing an alkyl group. Here, *any* amine and *any* carboxylic acid can decompose upon interaction with ionizing radiation via atomic hydrogen loss yielding the $RCHNH_2$ (R5) and RCHCOOH radicals (R6), respectively, which can then recombine with HOCO (R7) and NH_2 (R8), respectively, to form an amino acid. The amines and carboxylic acids can be synthesized from alkanes and carbon dioxide and ammonia, respectively, upon interaction with ionizing radiation as demonstrated experimentally (Kim & Kaiser 2010, 2011; Bennett & Kaiser 2007; Socrates 2001; R9–R12). Considering our experimental results, this reaction sequence can explain the formation of all non-polar amino acids with the carbon atoms in the hydrocarbon precursors classified as C1–C3: glycine (C1), alanine (C2), β-alanine (C2), and α-, β-, and γ-amino butyric acids (C3). The branched amino acids valine and leucine could be formed via branched hydrocarbons synthesized upon radiolysis of the C1–C3 precursors or after amino acid radiolysis. Further, the amino acids could also decompose upon interaction with ionizing radiation. This could either lead to carbon dioxide loss thus recycling the amine or to formation of an amino acid radical via hydroxyl loss (OH) (RI) and atomic hydrogen loss (RII). Both RI and RII can recombine in the matrix to the dipeptide.

4. ASTROPHYSICAL IMPLICATIONS

We have shown that the interaction of ionizing radiation in the form of energetic electrons with interstellar model ices at 10 K followed by annealing the processed ices to 300 K leads to the formation of nine amino acids and, most importantly, at least two dipeptides Gly–Gly and Leu–Ala. With the exception of nVal, the amino acids produced here were also detected in the Murchison meteorite. The amino acids can be grouped in two classes: proteinogenic amino acids (Gly, Ala, Leu, Val) and non-proteinogenic amino acids (β-Ala, α,β, and γ-amino butyric acid, nVal). The dipeptides (Gly–Gly, Leu–Ala) and possibly Ala–Gly and Ala–Ala that were identified in our experiments all contain proteinogenic amino acids.

It is important to place these findings in the context of molecular evolution on Earth. First, early functional peptides on Earth are suggested to have three to eight amino acids including Asp, Gly, Ala, and Val (Ehrenfreund et al. 2001; van der Gulik et al. 2009). Three of these amino acids (Gly, Ala and Val) were formed in our experiments and have also been identified in meteorites. Most importantly, our results rationalize the presence of the dipeptide Gly–Gly in the Murchison meteorite (Shimoyama & Ogasawara 2002). The shorter peptides could have acted as catalysts to yield longer peptides and proteins which gradually evolved to functional enzymes on early Earth (Ehrenfreund et al. 2001; van der Gulik et al. 2009). Second, since the stability of mineral-embedded peptides toward ionizing radiation has been established experimentally (Boillot et al. 2002), dipeptides could have survived on meteoritic parent bodies since the formation of our solar systems. We conclude that at least a fraction of the dipeptides on primitive Earth could have originated from meteorite or comet infall, thus supporting a potential exogenous source of prebiotic molecules on early Earth. This sequence of events presents an alternative to

Table 1
Infrared Absorption Features of the Residues at 300 K Formed upon Annealing the Irradiated Ice Mixtures

Experimental Absorption (cm^{-1})	Literature Absorption (cm^{-1})	Assignment
3236[a] (3200)[b]	3540–3270	N–H stretching
2964[a] (2965)[b]	2975–2950	$–CH_3$ asym. stretch
2952[a] (2945)[b]	2990–2930	$\nu(NH_4^+)$
2929[a] (2929)[b]	2940–2915	$–CH_2–$ asym. stretch
2870[a] (2869)[b]	2885–2865	$–CH_3$ sym. stretch
2769[a] (2780)[b]	2780–2530	$–NH_3^+$ sym. stretch
1667[a] (1687)[b]	1690–1670 (polar solids)[1]	Amide I
	1670–1650 (apolar phase)[1]	ν(C=O, carbonyl)
	1680–1630 (polar solids)[2]	[1]Primary amide
	1700–1665 (apolar phase)[2]	[2]Secondary amide
	1670–1630 (solid, apolar)[2]	[2]Tertiary amide
1590[a] (1611)[b]	1650–1620 (polar solids)[1]	Amide II
	1620–1590 (apolar phase)[1]	[1]Primary amide
	1570–1515 (polar solids)[2]	[2]Secondary amide
	1550–1510 (apolar phase)[2]	
	1620–1560	$\nu_{as}(–CO_2^-)$
1455[a] (1461)[b]	1465–1440	$–CH_3$ asym deform, $–CH_2–$ scissor
1382[a] (1381)[b]	1390 – 1370	$–CH_3$ sym deform
1371[a] (1387)[b]	1440 – 1335	$\nu_s(–CO_2^-)$
2400[c]	2403	N–D stretching
2221[c] (2219)[d]	2218	$–CD_2–/–CD_3$ stretching
2190[c]	2250–2150	$\nu(ND_4^+)$
2089[c]	2088	$–ND_3^+$
1719[c]	1730–1710	Carbonyl
1657[c] (1670)[d]	1680–1630	Amide I
1606[c] (1606)[d]	1650–1510	Amide II
1388[c] (1384)[d]	1440–1335	$\nu_s(–CO_2^-)$

Notes. Literature values were taken from Socrates (2001).
[a] Carbon dioxide (CO_2), ammonia (NH_3), and hydrocarbons methane, ethane, propane, *n*-butane, *n*-pentane, and *n*-hexane.
[b] Carbon dioxide (CO_2), ammonia (NH_3), and hydrocarbons methane, ethane, and propane.
[c] Carbon dioxide (CO_2), D-ammonia (ND_3), and perdeuterated hydrocarbons methane, ethane, and propane.
[d] Carbon dioxide (CO_2), ammonia (NH_3), and perdeuterated hydrocarbons methane, ethane, and propane.

competing theories like carbonyl sulfide (COS) mediated synthesis of polypeptides in aqueous media on early Earth (Leman et al. 2004) or a thermally triggered condensation of amino acids involving mineral catalysis at elevated temperatures reaching 473 K (Viedma 2000; Chyba & Hand 2005).

In conclusion, our experiments have established the feasibility that dipeptides—a key component in the assembly of proteins—can be formed in interstellar model ices abiotically at 10 K via ionizing radiation. Note that a recent re-analysis of the Murchison meteorite suggests the existence of high molecular weight organic molecules (Schmitt-Kopplin et al. 2010). Unfortunately, the mass spectra are too congested in the low mass range, so we cannot confirm that the dipeptides Ala–Ala and Leu–Ala have been also detected in Murchison meteorite. Our study will catalyze much needed successive investigations of the mechanism by which astrobiologically important molecules such as dipeptides and potentially higher polymers are formed on interstellar grains. However, no study can resolve all open questions simultaneously. Prospective work will investigate the role of water ice in the synthesis of dipeptides as water may act as a stabilizing energy transfer medium. Further, an

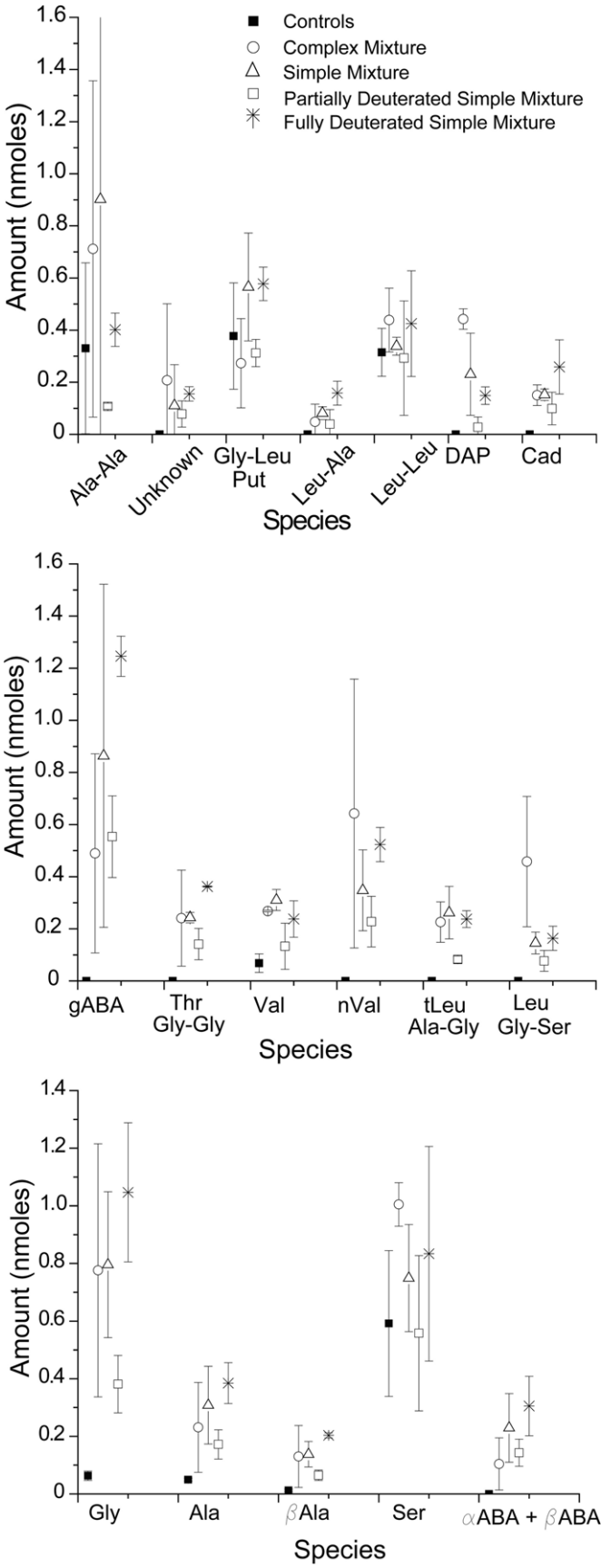

Figure 6. Species recovered in samples of interstellar ice analogs irradiated by simulated galactic cosmic radiation. The values and error bars are calculated based on two irradiation experimental repeats. The full procedural controls consisting of fully processed but non-irradiated targets are the controls presented in this figure.

Table 2
Control-corrected Masses (ng) of Amino Acids, Diamines, and Dipeptides Observed

Experiment	Full Procedural Control	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
		C1–C6, NH_3, CO_2		C1–C3, NH_3, CO_2		D–C1–C3, NH_3, CO_2		D–C1–C3, ND_3, CO_2	
Leu–Leu	8 ± 2	1	5	1	. . .[a]	. . .[a]	3	. . .[a]	6
Leu–Ala	0	0	2	1	2	0	2	3	4
Cad	0	1	2	1	2	1	1	2	3
Gly–Leu/ Put[b]	7 ± 4	. . .[a]	0.3	6	1	. . .[a]	. . .[a]	3	5
Put/Gly–Leu[b]	3 ± 2	. . .[a]	0.2	3	0.4	. . .[a]	. . .[a]	1	2
DAP	0	3	3	2	1	0.4	0	1	1
Ala–Ala	5 ± 5	. . .[a]	13	. . .[a]	24	. . .[a]	. . .[a]	2	0.4
Leu/Gly–Ser[b]	0	4	8	2	2	1	1	2	3
Gly–Ser/Leu[b]	0	5	10	2	3	1	2	2	3
tLeu/Ala–Gly[b]	0	2	4	3	4	1	1	3	3
Ala–Gly	0	3	4	3	5	1	1	3	4
nVal	0	3	12	3	5	2	3	6	7
Val	0.8 ± 0.4	2.4	2.3	3.2	2.5	0.02	1.5	1.4	2.6
Thr/Gly–Gly[b]	0	1	4	3	3	1	2	4	4
Gly–Gly/Thr[b]	0	1	5	3	4	1	2	5	5
γABA	0	2	8	4	14	5	7	13	12
αABA/βABA	0	4	17	15	32	11	18	24	39
Ser	60 ± 30	40	50	3	30	. . .[a]	20	. . .[a]	50
βAla	1.1 ± 0.4	3.7	17	8.4	14	3.5	5.8	16	18
Ala	4.4 ± 0.5	6.3	26	15	32	7.7	14	25	34
Gly	5 ± 1	30	80	40	70	20	30	60	90
Total dipeptide	8–20	1–10	20–50	2–20	30–40	0–4	5–10	4–20	10–30

Notes.
[a] Peak observed but not above blank levels.
[b] The two species indicated co-elute.

objective study on how dipeptide formation depends on distinct interstellar irradiation fields is warranted as is the role of potentially mineral-catalyzed formation of peptides on interstellar ices. Also, mechanisms for chiral selection such as template absorption, which might induce an enantiomer-rich synthesis, or an amino acid synthesis induced by circularly polarized ultraviolet light, shall be explored in future laboratory studies (Breslow & Cheng 2009). Finally, future studies should be performed to analyse the diastereomers of the dipeptides that are formed. Since the amino acids are formed as racemic mixtures, four distinct diastereomers of Leu–Ala and Ala–Ala should be seen. Unfortunately, diastereomeric analysis is beyond the scope and technical capabilities of the present project and would require a much larger amount of sample. Nevertheless, this study reveals unambiguously that extraterrestrial polypeptides, once produced, could have inseminated early Earth via comets and meteorites, and catalyzed its biological evolution.

We thank Professor P. B. Price (UC Berkeley, Physics Department) and Dr. Max Bernstein (NASA) for helpful comments and discussions and Dr. Thomas Chiesl (UC Berkeley) for his assistance in the initial phase of the project. Microdevices were fabricated in the UC Berkeley Bionanotechnology Center. This work was supported by the US National Science Foundation "Collaborative Research in Chemistry Program" (NSF-CRC; CHE-0627854). Research at Berkeley was supported by the Mathies Royalty Fund.

APPENDIX

Discussion of Peak Assignments by μCE: The area of each peak was determined using PeakFit software and then used to calculate the concentration of each species based on a standard curve. Figure 6 displays the equivalent molar amounts of each species produced in each experiment without any blank or control correction. Two types of controls were performed: a μCE blank and a full procedural control where only the irradiation of the target was omitted. The μCE blank was labeled alongside every sample and analyzed both directly before and directly after the samples of interest. Only five species were seen in the μCE blank: Leu–Leu, Gly–Leu/Put, and Ala were observed at consistent levels with daily variations of less than 10%; Ser (<3 nM) and Gly (<50 nM) had variable contributions. The μCE blanks contained such low concentrations of species compared to the full procedural controls that only the full procedural control concentrations were utilized for data processing. Concentrations corrected for the full procedural controls were used to calculate the mass of each species recovered from the targets, and these masses are presented in Table 2. The observed mass of each species in the procedural control is also listed for comparison in the left column of Table 2. The upper bound of each species is given for peaks representing potentially co-eluting pairs (Gly–Leu/Put, Leu/Gly–Ser, Leu/Ala–Gly, Thr/Gly–Gly). The total amount of dipeptides observed is listed in the bottom row, and ranges from 0–50 ng.

In each sample, there is a clear trend of simpler species being generated at higher levels than more complex species. Glycine is always produced in the highest proportion ($\sim$ 40%), followed by alanine (10%–20%) and β-alanine (6%–9%). This is expected due to the ratio of hydrocarbons used, and this general trend can be used to make arguments about relative concentrations of co-eluting species. In these experiments, we consistently see the enrichment of straight-chain species (Ala, nVal). This trend is best explained by the fact that straight-chain hydrocarbons were

used in the irradiation experiments as carbon-bearing precursor molecules.

Some analysis is required to determine which observed dipeptides are unambiguously produced as a result of ice irradiation. First, the variability and relatively high levels of the dipeptide Ala–Ala in the full procedural blanks makes it impossible to unambiguously confirm its production as a result of ice irradiation (Figure 6). Second, the dipeptides Gly–Gly and Ala–Gly require careful analysis because they co-elute with the amino acids Thr and tLeu, respectively. The relative amounts of all isomers of aminobutyric acid are enriched in the simple hydrocarbon mixture samples compared to the complex mixture with C4-C6 hydrocarbons (Figure 6). A similar trend would be expected for the Thr/Gly–Gly peak if it were due solely to Thr. However, the relative amount of the peak migrating as Thr/Gly–Gly remains constant across the various samples; this observation is consistent with the interpretation of this peak as being due to Gly–Gly instead of Thr. Additionally, we see relatively high levels (similar to those observed for Val) of the peaks due to tLeu/Ala–Gly in all samples (Figure 6). This is inconsistent with the interpretation of these peaks as being due solely to the amino acid because the general trend observed is preferential production of amino acids with fewer carbon species, and Val (5C) is smaller than tLeu (6C). Therefore, there must be a significant contribution to the peak labeled tLeu/Ala–Gly from the Ala–Gly dipeptide. Thirdly, Ser levels are significant and highly variable in both controls and real samples (Figure 6), masking our ability to identify its irradiation-associated production. We do see a high level of the peak labeled as Leu/Gly–Ser that exhibits an intensity similar to that of nVal. Because Leu is longer than nVal, this observation is inconsistent with the concept that the Leu/Gly–Ser peak is due solely to Leu. Therefore, the presence of some Gly–Ser in the samples is not inconsistent with these data. We conclude that the Leu/Gly–Ser peak cannot be fully attributed to either Leu or Gly–Ser and is most likely a combination of the two species. Peaks migrating at the expected positions of Leu-containing dipeptides (Leu–Leu, Leu–Ala) are observed. However, Leu–Leu is not produced at levels significantly above the procedural controls, and the Leu–Ala peak is not reproducibly produced in detectable levels in all samples. Finally, we expect very low levels of dipeptides made up of the lower concentration amino acids (Thr, Val, nVal, Leu, tLeu). Dipeptides that contain an undetected amino acid would be very unlikely, and our results are consistent with this expectation. The remaining unassigned peaks may be due to ABA dipeptides which are not commercially available for comparison.

REFERENCES

Bennett, C. J., Jamieson, C., Mebel, A. M., & Kaiser, R. I. 2004, PCCP, 6, 735
Bennett, C. J., & Kaiser, R. I. 2007, ApJ, 660, 1289
Bernstein, M. P., Dworkin, J. P., Sandford, S. A., Cooper, G. W., & Allamandola, L. J. 2002, Natur, 416, 401
Boillot, F., Chabin, A., Buré, C., et al. 2002, OLEB, 32, 359
Breslow, R., & Cheng, Z.-L. 2009, PNAS, 106, 9144
Charnley, S. B., Ehrenfreund, P., & Kuan, Y.-J. 2001, AcSpA, 57, 685
Chyba, C. F., & Hand, K. P. 2005, ARA&A, 43, 31
Drouin, D., Couture, A. R., Gauvin, R., et al. 2001, Monte Carlo Simulation of Electron Trajectory in Solids (CASINO VERSION. 2.42; Sherbrooke: Univ. Sherbrooke)
Ehrenfreund, P., Glavin, D. P., Botta, O., Cooper, G., & Bada, J. L. 2001, PNAS, 98, 2138
Elsila, J. E., Glavin, D. P., & Dworkin, J. P. 2009, M&PS, 44, 1323
Gibb, E. L., Whittet, D. C. B., Boogert, A. C. A., & Tielens, A. G. G. M. 2004, ApJS, 151, 35
Holtom, P. D., Bennett, C. J., Osamura, Y., Mason, N. J., & Kaiser, R. I. 2005, ApJ, 626, 940
Kaiser, R. I., & Roessler, K. 1997, ApJ, 475, 144
Kaiser, R. I., & Roessler, K. 1998, ApJ, 503, 959
Kim, Y. S., & Kaiser, R. I. 2010, ApJ, 725, 1002
Kim, Y. S., & Kaiser, R. I. 2011, ApJ, 729, 68
Lafosse, A., Bertin, M. M., Domaracka, A. A., et al. 2006, PCCP, 8, 5564
Leman, L., Orgel, L., & Ghadiri, M. R. 2004, Sci, 306, 283
Meinert, C., Filippi, J.-J., de Marcellus, P., d'Hendecourt, L., & Meierhenrich, U. J. 2012, ChemPlusChem, 77, 186
Muñoz Caro, G. M., Meierhenrich, U. J., Schutte, W. A., et al. 2002, Natur, 416, 403
Navarro-González, R., Vargas, E., de la Rosa, J., Raga, A. C., & McKay, C. P. 2010, JGR., 115, E12010
Oró, J. 1961, Natur, 190, 389
Remusat, L., Guan, Y., Wang, Y., & Eiler, J. M. 2010, ApJ, 713, 1048
Schmitt-Kopplin, P., Gabelica, Z., Gougeon, R. D., et al. 2010, PNAS, 107, 2763
Shimoyama, A., & Ogasawara, R. 2002, OLEB, 32, 165
Skelley, A. M., Scherer, J. R., Aubrey, A. D., et al. 2005, PNAS, 102, 1041
Socrates, G. 2001, Infrared and Raman Characteristic Group Frequencies (3rd ed.; New York: Wiley)
Stano, P., & Luisi, P. L. 2009, in Prebiotic Evolution and Astrobiology, ed. A. Lazcano & J. T. Wong (Austin, TX: Landes Bioscience), 94
Stockton, A. M., Chiesl, T. N., Lowenstein, T. K., et al. 2009, AsBio, 9, 823
van der Gulik, P., Massar, S., Gilis, D., Buhrman, H., & Rooman, M. 2009, J. Theoret. Biol., 261, 531
Viedma, C. 2000, OLEB, 30, 549
Weber, A. L., & Pizzarello, S. 2006, PNAS, 103, 12713