



**HAL**  
open science

# Gas chromatography coupled-to Fourier transform orbitrap mass spectrometer for enantioselective amino acid analyses: Application to pre-cometary organic analog

A. Garcia, C. Serra, Q. Blancart Remaury, A.D. Garcia, M. Righezza, Cornelia Meinert, P. Poinot, G. Danger

## ► To cite this version:

A. Garcia, C. Serra, Q. Blancart Remaury, A.D. Garcia, M. Righezza, et al.. Gas chromatography coupled-to Fourier transform orbitrap mass spectrometer for enantioselective amino acid analyses: Application to pre-cometary organic analog. *Journal of Chromatography A*, 2023, 1704, pp.464118. <10.1016/j.chroma.2023.464118>. <hal-04150942>

**HAL Id: hal-04150942**

**<https://hal.science/hal-04150942v1>**

Submitted on 4 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1 **Gas chromatography coupled-to Fourier transform orbitrap mass spectrometer for enantioselective**  
2 **amino acid analyses: application to pre-cometary organic analog**

3 A. Garcia<sup>1</sup>, C. Serra<sup>1,2</sup>, Q. Blancart Remaury<sup>2</sup>, A. D. Garcia<sup>3</sup>, M. Righezza<sup>1</sup>, C. Meinert<sup>3</sup>, P. Poinot<sup>2</sup>, G.  
4 Danger<sup>1,4\*</sup>

5 <sup>1</sup> Aix-Marseille Université, CNRS, Institut Origines, Laboratoire PIIM, Marseille, France

6 <sup>2</sup> UMR CNRS 7285, Institut de Chimie des Milieux et Matériaux de Poitiers (IC2MP), E.BiCoM Team,  
7 University of Poitiers, 4 rue Michel-Brunet, TSA 51106, 86073, Poitiers, Cedex 9, France

8 <sup>3</sup> Université Côte d'Azur, Institut de Chimie de Nice, UMR 7272 CNRS, F-06108 Nice, France

9 <sup>4</sup> Institut Universitaire de France (IUF)

10

11 **ABSTRACT:**

12 Gas chromatography (GC) is a separation technique commonly developed for targeted *in situ* analyses in  
13 planetary space missions. It is coupled with low-resolution mass spectrometry to obtain additional structural  
14 information and allow compound identification. However, ground-based analyses of extraterrestrial samples  
15 have shown the presence of large molecular diversities. For future targeted *in situ* analyses, it is therefore  
16 essential to develop new technologies. High resolution mass spectrometry (HRMS) is currently being spatialized  
17 using FT-orbitrap-MS technology. In this contribution, the coupling of gas chromatography with FT-orbitrap-  
18 MS is studied for targeted amino acid analyses. The method for enantioselective separation of amino acids was  
19 optimized on a standard mixture comprising 47 amino acid enantiomers. Different ionization modes were  
20 optimized, chemical ionization with three different reactive gases (NH<sub>3</sub>, CH<sub>4</sub> and NH<sub>3</sub>/CH<sub>4</sub>) and electron impact  
21 ionization at different electron energies. Single ion and full scan monitoring modes were compared, and  
22 detection and quantification limits were estimated by internal calibration under the optimized conditions. The  
23 GC-FT-orbitrap-MS demonstrated its ability to separate 47 amino acid enantiomers with minimal co-elution.  
24 Furthermore, due to the high mass resolution and accuracy of FT-orbitrap-MS, with mass extraction, the *S/N* is  
25 close to zero, allowing average LOD values of 10<sup>-7</sup>M, orders of magnitude lower than conventional GC-MS  
26 techniques. Finally, these conditions were tested for enantioselective analysis of amino acids on an analog of a  
27 pre-cometary organic material showing similarities to that of extraterrestrial materials.

28 Keywords: amino acids, GC-FT-Orbitrap-MS, gas chromatography, high-resolution mass spectrometry,  
29 cometary ice analogs.

30

31 **1. INTRODUCTION**

32 The understanding of the origin of the solar system is partly linked to the characterization of the organic content  
33 of comets or asteroids. These interplanetary bodies are indeed considered to have evolved little since the  
34 formation of the solar system and thus possess a primitive organic content<sup>1-3</sup>. To characterize this organic  
35 matter, ground-based or space observations are possible, but are limited to the identification of molecules either  
36 in the gas phase<sup>4-6</sup> or to the surface characterization mainly by infrared spectroscopy<sup>7-9</sup>. However, space  
37 missions have been designed to visit interplanetary bodies in order to obtain a more precise characterization of  
38 their organic content<sup>10</sup>. In several of these missions, the basic components of biomolecules, such as amino acids,  
39 are searched for. For example, the Rosetta mission analyzed a cometary environment *in situ* and detected various  
40 organic molecules such as the simplest amino acid: glycine, using mass spectrometers<sup>11</sup>. However, as meteorite  
41 analyses revealed<sup>12</sup>, the molecular content of these bodies is much more diverse than what is suggested by these  
42 *in situ* analyses. Indeed, in some meteorites, up to 90 different amino acids have been observed<sup>13</sup>. It is therefore  
43 necessary to develop alternative technologies allowing an unambiguous identification of the targeted molecules.

44 In this context, the National Centre for Space Studies (CNES) is transferring the orbitrap technology for space  
45 mission, the cosmorbitrap<sup>14,15</sup>. Indeed, high resolution mass spectrometry has shown its ability to decipher the  
46 impressive molecular diversity contained in the soluble or insoluble organic fractions of meteorites<sup>12,16-21</sup>.  
47 Furthermore, gas chromatography (GC) is currently used as a separation technic for *in situ* analyses, and was on  
48 board of various space missions usually coupled with low resolution mass spectrometry<sup>22-26</sup>. Since 2015, the

49 FT-Orbitrap mass spectrometer has been coupled to gas chromatography via an electron impact ionization source  
50 <sup>27,28</sup>. This new technology could thus be an alternative to the GC-MS systems of previous space missions.  
51 However, before developing such a coupling for space exploration, it is important to define the potential interest  
52 of this technology compared to previous systems. GC-FT-Orbitrap-MS was previously used to analyze volatile  
53 organic compounds generated by the weathering of pre-cometary ice analogs <sup>29</sup>. This work demonstrated the  
54 importance of the high resolution of the mass spectrometer in compound identification. Indeed, in the laboratory,  
55 it is possible to experimentally simulate the evolution of a cometary ice, which provides molecular compounds  
56 that are considered as analogs of those present in cometary environments <sup>30</sup>. These experimental simulations  
57 allow to test the efficiency of the analytical methodologies developed for space analytical tools, and to estimate  
58 the interest to develop new analytical tools for space missions as in the case of the coupling between a gas  
59 chromatography and a FT-Orbitrap-MS.

60 In this contribution, we further investigated the potential of a GC-FT-Orbitrap-MS, determining the optimal  
61 parameters for targeted amino acid analyses. Using amino acid standard mixtures, the resolution was first  
62 optimized to obtain optimal enantioseparation. We then determined the optimal electron impact and chemical  
63 ionization parameters to obtain the best ionization yields. Moreover, the sensitivity of the GC-FT-Orbitrap-MS  
64 was defined using an internal calibration. The optimal conditions were finally applied to the search for amino  
65 acids in an organic residue analog to pre-cometary organic matter. The results obtained were compared to  
66 previous analyses of amino acids in similar samples, as well as to analytical instruments using the same  
67 separation technology, GC × GC-TOF-MS, and liquid chromatography coupled to FT-Orbitrap-MS.

## 68 2. MATERIAL AND METHODS

### 69 2.1 Chemicals and solutions

70 For amino acid analyses, each amino acid was prepared individually and then mixed together in 0.1 M  
71 hydrochloric acid (HCl) (**required for derivatization**) to obtain a stock solution of  $10^{-4}$  M (Table 1). Serial  
72 dilutions were prepared for calibration curves. The 0.1 M HCl solution was prepared by diluting 6 M HCl (for  
73 amino acid analysis ampoule, Merck) in ultra-pure water produced by a Direct-Q® 3 UV water purification  
74 system. All amino acids and chemicals used were from Sigma-Aldrich, Fluka, or Acros Organics.

### 75 2.2 Derivatization procedure

76 **For GC analyses only, derivatization was performed to make the amino acids volatile, and this method of**  
77 **sample preparation is already used for space missions, demonstrating its feasibility** <sup>23,24</sup>. Amino acid  
78 solutions were derivatized into their *N* (*O*, *S*)-ethoxy-carbonylheptafluorobutylester (ECHFBE) derivatives  
79 according to the protocol developed by Meinert *et al.* <sup>31</sup>. **We choose this protocol because it was optimized for**  
80  **$\alpha$ -alkyl and  $\alpha$ -dialkyl amino acid** <sup>29</sup>. In a conical reaction flask (Reactivial, Thermo Scientific™), a 50  $\mu$ L  
81 volume of aqueous amino acid solution in 0.1 M HCl (**derivatization is done in acid condition**) reacted with 19  
82  $\mu$ L of 2,2,3,3,4,4,4-heptafluoro-1-butanol and 6  $\mu$ L of pyridine. After 15 s of stirring, 5  $\mu$ L of ethyl  
83 chloroformate were added and the resulting solution stirred vigorously for 15 seconds. The ECHFBE derivatives  
84 were extracted with a  $10^{-5}$  M methyl laurate (internal standard) chloroform solution. Only, the organic phase  
85 containing the ECHFBE derivatives was then transferred into 1 mL GC vials equipped with 100  $\mu$ L inserts for  
86 GC-FT-Orbitrap-MS analysis. Note that this derivatization method does not provide identical yields among  
87 different amino acid groups, especially discriminating  $\alpha$ ,  $\alpha$ -dialkyl amino acids as well as  $\gamma$ -amino acids <sup>32</sup>. **This**  
88 **derivatization method does not generate a racemization process** <sup>28,29</sup>, **verified with the pure D and L**  
89 **enantiomers of 2-ABA.**

### 90 2.3 Synthesis of organic residues analog to pre-cometary organic material

91 A pre-cometary organic analog was formed from an ice including H<sub>2</sub>O, <sup>13</sup>CH<sub>3</sub>OH and NH<sub>3</sub> in proportion of  
92 3:1:1. The corresponding gas mixture was deposited in a stainless steel chamber on a cold finger at low pressure  
93 ( $10^{-7}$  to  $10^{-8}$  mbar) and low temperature (77 K) forming an ice analog to the ones observed in dense molecular  
94 clouds on silicate grains. The ice formation was concomitant to its irradiation with a dihydrogen UV lamp  
95 (mainly emitting at 121 nm) to simulate stellar radiation. After 72 h of deposition and simultaneous irradiation,  
96 the photoprocessed ice was slowly heated up to 300 K to obtain an analog of pre-cometary organic matter. To be  
97 analyzed, the residue was recovered in ultrapure MilliQ water. The solution was then dried and recovered in 0.1  
98 M HCl to be hydrolyzed during 24 h at 110 °C. <sup>13</sup>C-labeled methanol is used to overcome the problem related to  
99 potential contamination with natural <sup>12</sup>C-amino acid.

## 100 2.4 GC-FT-Orbitrap-MS configuration

101 Analyses were performed by a Trace 1310 gas chromatograph (GC) coupled to a Q-Exactive Orbitrap™ mass  
102 spectrometer (MS) from Thermo Fisher Scientific. Injections were performed with an auto-sampler (AI 1310  
103 from Thermo Fisher) in splitless mode (splitless time: 1 min) with an injector temperature of 230 °C. Helium  
104 was used as carrier gas with a flow rate of 1 mL min<sup>-1</sup> and a purge rate of 5.0 mL min<sup>-1</sup>. The GC injector has  
105 been modified to receive a pyrolysis system (Multi-Short Pyrolyzer EGA/PY3030 D from Frontier Lab) that can  
106 only use split liners. Verifications were performed to validate the absence of potential impacts on our operational  
107 conditions using a split liner for splitless analysis. Amino acids were separated on two Chirasil-L-Val columns  
108 (for each column 25 m x 0.25 mm x 0.12 µm film thickness, Agilent) connected with a Valco connector. The  
109 duration of the oven temperature program was 90 min with a solvent delay of 14 min. The optimized temperature  
110 program was as follows: 40 °C for 1 min, then increased to 80 °C with a slope of 10 °C min<sup>-1</sup> during 10 min  
111 then 2 °C min<sup>-1</sup> to reach 190 °C with an isotherm during 20 min. The transfer line was set at 250 °C. The *m/z*  
112 range was 50–400 with a FWHM resolution fixed at 60 000, a target AGC value at 10<sup>6</sup> and a max IT at 200 ms.  
113 Two ionization modes have been tested and optimized with electron energy at 70 eV: electron impact and  
114 chemical ionization. For the latter, three different reactant gases were tested in positive and negative mode,  
115 ammonia, methane and ammonia-methane mixture (10/90), with an ionization source at 300 °C and a reactant  
116 gas flow rate of 1.5 mL min<sup>-1</sup> in the ion volume. For the electron impact ionization, the ionization source  
117 temperature was set at 250 °C. Analyses in selected ion monitoring (SIM) were performed with the Orbitrap cell  
118 at a FWHM resolution of 60 000, an optimized target AGC of 5.10<sup>4</sup>, a max IT of 200 ms and a mass to charge  
119 isolation window of 0.4. **In order to optimize the GC method for a good compromise between resolution  
120 and analysis time, some parameters were optimized. The choice of the analytical column chirasil L-Val  
121 was made and justified by Meinert et al<sup>29</sup>. The length of the column as well as the temperature gradient  
122 were optimized in order to obtain a resolution of the amino acids and their enantiomers while minimizing  
123 the analysis time.**

## 124 2.5 GC × GC–TOF-MS configuration

125 The enantioselective GC × GC–TOF-MS analyses<sup>32,33</sup> were carried out using a Pegasus IV D instrument  
126 coupled to a reflectron TOF-MS (LECO corp.) operating at 100 Hz storage rate, with a 50–400 amu mass range,  
127 a microchannel plate detector voltage of 1.8 kV, and a solvent delay of 12 min. Sample aliquots of 1 µL were  
128 injected with an automatic liquid sampler (Agilent 7693A) in splitless mode. The injector and ion source  
129 temperatures were kept at 230 °C and the transfer line at 240 °C. The column set consisted of a Chirasil-L-Val  
130 primary column (25 m length × 0.25 mm inner diameter, 0.12 µm film thickness, Agilent-Varian, Santa Clara,  
131 California, US) coupled to a DB Wax secondary column (1.6 m length × 0.25 mm inner diameter, 0.1 µm film  
132 thickness, Agilent) via a siltite µ-union. Helium was used as a carrier gas at a constant flow of 1 mL min<sup>-1</sup>. The  
133 primary oven was operated as follows: 40 °C for 1 min, 10 °C.min<sup>-1</sup> to 80 °C (10 min hold time), and 2 °C.min<sup>-1</sup>  
134 to 190 °C (10 min hold time). The secondary oven operated with a constant positive temperature offset of 35 °C,  
135 the thermal modulator hot jets temperature offset was set at 15 °C and the modulation period was 6 s. The data  
136 were processed using Leco corp. ChromaTOF® software.

## 137 2.6 LC-FT-Orbitrap-MS configuration

138 Accela Ultra-Performance Liquid Chromatography (UPLC) pumps were interfaced with a High-Resolution Mass  
139 Spectrometer (Q-Exactive, Hybrid Quadrupole-Orbitrap Mass Spectrometer, Thermo Fisher Scientific, Waltham,  
140 MA, USA). Amino acids enantiomers were separated on a chiral Crownpak CR (+) column (150 mm × 4.0 mm,  
141 5 µm, Chiral Technologies, Europe). A guard column Crownpak CR (10 mm × 4.0 mm, 5 µm, Chiral  
142 Technologies) was fitted to the analytical column. Elution was carried out at 600 µL min<sup>-1</sup> with water:methanol  
143 (95:5) with 0.5% trifluoroacetic acid (Sigma-Aldrich) during 20 min. The UPLC oven temperature was held at  
144 30 °C. Compound ionization was performed by heated electrospray ionization (HESI) in positive mode. Mass  
145 spectrometer capillary voltage was set at 2.8 kV, the capillary temperature was held at 400 °C and aux gas heater  
146 temperature at 300 °C. The desolvation line temperature was held at 250 °C. Sheath gas flow, aux gas flow and  
147 sweep gas flow were set at 40, 30 and 0 u.a. respectively<sup>34</sup>.

## 148 2.7 Amino acid database generation and calibration curves for GC-FT-Orbitrap-MS

149 Analyses were performed using the Xcalibur software (Thermo Scientific). The data obtained in TIC were  
150 processed with Qual Browser Xcalibure (Thermo Scientific). For each amino acid, mass extraction was

151 performed on the TIC data. A GC-FT-Orbitrap-MS database was obtained by injecting each **achiral amino acid**  
152 **(e.g. glycine, sarcosine or  $\beta$ -alanine)** as well as **racemic mixture for chiral amino acids individually** (except  
153 for DL-2-ABA for which the enantiomers were injected individually) at a concentration of  $10^{-4}$  M, to  
154 determine their retention times and fragmentation pattern. Thanks to the database, a SIM method was edited with  
155 a characteristic ion for each amino acid and a retention time window with a minimum of one minute around their  
156 respective retention time (Table S1). In post analyses, a Gaussian type smoothing was applied to compensate for  
157 the overlap of several masses in the same retention time window. This database established the probability of  
158 correlation with the experimental spectra based on the retention time and the fragmentation profile using the 30  
159 most intense  $m/z$  for each amino acid.

160 Quantification was obtained by integrating the characteristic ion of each amino acid (AA) divided by the peak  
161 integration of the internal standard (AIS), AA/AIS. The determination of LOD, LOQ and linearity of the  
162 calibration curves were obtained by the NeoLiCy® software (VWR). This software is designed for statistical  
163 evaluation applied to any type of analytical method. It includes ready-to-use models for the configuration of  
164 statistical tests and their associated calculations. **The LOD and LOQ are calculated as follows:  $LOD = 3 \times \sigma / S$ , and  $LOQ = 10 \times \sigma / S$** , where S is the slope of the calibration curve and  $\sigma$  is the standard deviation of the  
165 response. Calibration curves were performed with seven concentrations ranging from  $5 \times 10^{-7}$  to  $1 \times 10^{-6}$  M,  
166 in triplicate of sample preparation and duplicate of injection. Moreover, the data processed by NeoLiCy  
167 indicated that the linearity is no longer verified beyond four range points. It was thus chosen to work on the four  
168 lowest concentrations based on the sensitivity of each amino acid (Table S2).  
169

## 170 3. RESULTS

### 171 3.1 Chromatographic separation, high resolution, and mass accuracy

172 The implemented chromatographic method allowed for resolving almost all forty-seven amino acids (Figure 1,  
173 and Table 2), with only three coelutions being observed. Co-elution concerns L-norleucine/L- $\beta$ -leucine, D-  
174 methionine/D-glutamic acid, and L-methionine/L-glutamic acid. However, because of the high resolution of the  
175 mass spectrometer, it is possible to discern these co-eluting analytes (Table 2). The high mass accuracy provided  
176 by the Orbitrap technology allows for the unambiguous identification and quantification of each amino acid due  
177 to the precision of the mass to charge ratio ( $m/z$ ) and, consequently, an average error of less than 2 ppm in the  
178 stoichiometric formula assignment (Table 2). Moreover, this precision in mass allows to greatly reduce the signal  
179 to noise ratio when working by mass extraction (Figure 1), with a measured noise of  $6 \times 10^5$  (absolute intensity)  
180 by mass extraction compared with  $1 \times 10^8$  in full scan.

### 181 3.2 Optimization of the ionization source parameters for optimal chemical or electron impact ionizations

182 In addition to the fragmentation pattern obtained with the EI ionization mode that directly informs on the  
183 chemical structure of an unknown molecule, the detection of the molecular ion can support analyte identification,  
184 especially for compounds absent in the database. In this section, the chemical and electron impact ionizations  
185 were optimized to define the best parameters providing molecular ion detection for several amino acids. Electron  
186 impact was tested by modifying the electron energy. Chemical ionization was used with several reactant gases in  
187 positive and negative modes.

#### 188 3.1.2 Ionization by electronic impact (EI)

189 At 70 eV amino acid fragmentation is privileged and hence molecular ions are not detected for most of the amino  
190 acids, and if detected represent only minor peaks in the spectra. To verify the possibility of observing the  
191 molecular ion using lower EI ionization, three ionization energies were tested: 50 eV, 30 eV and 10 eV. Figure 2  
192 shows the results obtained for two amino acids, sarcosine, which is representative of the behavior of the set of  
193 amino acids, and cysteine, which shows a particular behavior. Even at these electron energies, the molecular ions  
194 were not observed (Figure 2). Fragmentation remained essentially the same, with a considerable decrease in  
195 fragment intensity with the decrease of the electron energy. Consequently, even at lower electron energy, the EI  
196 ionization mode is not valuable to obtain the molecular ion of amino acids.

#### 197 3.1.3 Ionization by chemical ionization (CI)

198 Chemical ionization, as a softer ionization mode than EI, ionizes first a reactant gas that then ionizes the target  
199 molecules. The compounds of interest undergo proton transfer (positive mode, PCI) or electron removal

200 (negative mode, NCI). Three types of reactive gases were tested in positive and negative mode: ammonia,  
201 methane and a mixture of ammonia/methane (90/10).

202 To determine the optimal CI mode for molecular ion detection, area and peak intensities of the molecular ions  
203 were obtained by their mass extraction. Our results indicate that the highest intensities are obtained with PCI  
204  $\text{NH}_3/\text{CH}_4$  (Table 3). However, for  $\alpha$ -amino acids, similar sensitivities are obtained with NCI  $\text{NH}_3$ , with the  
205 exception of DL-proline, probably due to its cyclic secondary amine (Table 3). In NCI, the other reactant gases  
206 do not show improved results. Moreover, independent of the ionization gas used in negative mode, less  
207 satisfactory results were obtained because of increased peak tailing and the constant presence of the  $m/z$  177,  
208 whose origin has not been identified yet, but biases the results.

209 Table 3 shows that the molecular ion is generally the most intense ion in the spectra in the  $\text{NH}_3$  and  $\text{NH}_3/\text{CH}_4$   
210 PCIs.  $\text{NH}_3$  provides additional information for the validation of amino acid identification by the presence of the  
211  $[(\text{M}+\text{H})+\text{NH}_3]^+$  adduct, which is generally the second most intense ion. Only cysteine deviates from the overall  
212 behavior because in none of the NCI or PCI modes the molecular ion is the peak base. Furthermore, with this  
213 amino acid the highest intensity of the molecular ion is observed in EI. This deviation from other amino acids is  
214 certainly due to the presence of the thiol (SH) function.

215 Overall, the EI mode is not effective for the search of amino acids by their molecular ions. Only the PCI  $\text{NH}_3$   
216 and  $\text{NH}_3/\text{CH}_4$  modes allow for an identification of all amino acids by their molecular ion, with the PCI  $\text{NH}_3$   
217 mode presenting an additional advantage that is the  $\text{NH}_3$  adduct. Therefore, we have chosen the PCI  $\text{NH}_3$  mode  
218 for further method optimization including the ionization source temperature as well as the flux of ionizing gas  
219  $\text{NH}_3$ . We found the optimized ionizing gas flow to be between  $1.5 \text{ mL min}^{-1}$  and  $2 \text{ mL min}^{-1}$  and the  
220 optimized ion source temperature to be  $300 \text{ }^\circ\text{C}$ . Nevertheless, the overall sensitivity of the EI mode is higher  
221 than any of the CI modes tested, independent of the reactant gas/mixture used (Table S4). As the fragmentation  
222 pattern obtained in EI mode offers valuable information for compound identifications, we thus focus our method  
223 development on the EI mode at  $70 \text{ eV}$  for the rest of the study.

### 224 3.3 Limits of detection and quantification using EI ionization mode at $70 \text{ eV}$

225 To evaluate the detectability of amino acids with a GC-FT-Orbitrap-MS, the limits of detection (LOD) and  
226 quantification (LOQ) were determined. The mass accuracy provided data almost free of noise. The estimations  
227 of the LOD and LOQ were thus not calculated according to the usual signal-to-noise ratio method but based on  
228 the calibration method using the NeoLiCy® software (see Materials and Methods).

229 Due to the difference in derivatization yields, the LOQ and LOD values vary by one order of magnitude among  
230 the amino acids. A ranking was made to classify the amino acids based on their detection threshold into lowest,  
231 intermediate and uppermost ones. Thirteen amino acids present an average LOD of  $1.98 \times 10^{-7} \text{ M}$ , with  
232 sarcosine presenting the lowest detection limit of  $1.50 \times 10^{-7} \text{ M}$ . The least detectable amino acids present an  
233 average LOD of  $2.18 \times 10^{-6} \text{ M}$  (Table 4). These values are one to two orders better than amino acids analyzed  
234 by other GC-MS techniques<sup>34,35</sup>, and only one to two orders of magnitude above GC $\times$ GC-MS analyses<sup>31,33</sup>.

### 235 3.4 Single ion monitoring vs. full scan monitoring

236 To compare the SIM acquisition mode with the full scan mode followed by mass extraction data processing, the  
237 same mixture of amino acids at the same concentration was injected successively in full scan and in SIM. Data  
238 were then treated as the area of the characteristic fragment ion over the area of the  $m/z$  87.04401 ion of the  
239 internal standard methyl laurate at  $10^{-5} \text{ M}$ . Compared to the full scan, SIM facilitates the search for amino acids  
240 with high LOD values as well as those exhibiting similar fragmentation profiles and retention times, as it is the  
241 case for GABA and 5-AVA in our mixture. However, the overlapping retention time windows in SIM mode  
242 (Table S1) due to co-elution or low resolution for some amino acids meant that the peaks did not have a mass  
243 specific to only one amino acid and thus could not be integrated directly. This problem was solved by working in  
244 mass extraction.

245 The comparison of SIM results to those of the full scan mode shows that due to the targeted analysis including  
246 derivatization and extraction procedures greatly minimizing any matrix effects, the advantage of SIM over the  
247 full scan mode is not decisive. Furthermore, because of the high resolution, the accuracy of the extracted mass in  
248 full scan mode limits the noise intensity. Therefore, the contribution of the SIM acquisition mode is

249 inconclusive, as the  $A_x/A_{IS}$  ratio in mass extraction is similar or even significantly better in full scan acquisition  
250 (Table 5).

251 Consequently, for analyses of samples including unknown amino acids, we found that the optimal conditions for  
252 a GC-FT-Orbitrap-MS method are (i) the EI as ionization mode at 70 eV in (ii) full scan with mass extraction of  
253 the base peak. If necessary, especially if the detected compounds are absent in the database, the PCI-NH<sub>3</sub> mode  
254 can be used to obtain information on the molecular ions, if the sensitivity is sufficient. Finally, the NCI- NH<sub>3</sub>  
255 mode can directly inform on the presence of  $\alpha$ -amino acids.

## 256 4. DISCUSSION

### 257 4.1 Application of optimized conditions to the detection of amino acids in a pre-cometary analog

258 A pre-cometary analog was formed at low temperature (77 K) and pressure ( $10^{-7}$  mbar) from an ice analog  
259 containing H<sub>2</sub>O:<sup>13</sup>CH<sub>3</sub>OH:NH<sub>3</sub> in ratio 3:1:1 irradiated at 121 nm during 72 h. The photo-processed ice was then  
260 heated up to 300 K to obtain an organic residue, which is considered as an analog to pre-cometary material. The  
261 resulting organic residue was hydrolyzed during 24 h at 100 °C in 6N HCl, and amino acids were searched for in  
262 the resulting sample. The analysis of the residue was carried out with the GC-FT-Orbitrap-MS in full scan with  
263 the EI mode at 70 eV after carrying out the ECHFBE derivatization. Data were then treated by mass extraction.  
264 The theoretical masses of the characteristic amino acid mass fragments were calculated by considering <sup>13</sup>C  
265 isotopes related to the amino acid backbone and the respective <sup>12</sup>C derivatization chemical groups (Table 6). The  
266 amino acid identifications in the residue were based on their retention times as well as their fragmentation  
267 profiles compared with those established in the amino acid database (Table S3).

268 Among the numerous amino acids present in such ice analog residues, 32 amino acid enantiomers were  
269 identified in this study using the database established with GC-FT-Orbitrap-MS (Table 6 and Figure 3). Among  
270 these, the proteinogenic amino acids such as alanine, glycine, proline, serine, valine and aspartic acid were  
271 identified as well as some non-proteinogenic ones. Chiral amino acids are detected as racemic mixtures. Amino  
272 acid quantification was performed to determine the abundances of each identified amino acid (Table 6). Glycine  
273 and alanine are the most abundant amino acids, with concentration of  $5.36 \times 10^{-4}$  M and  $9.30 \times 10^{-5}$  M,  
274 respectively. Furthermore, amino acid concentrations decrease with the increase in molecular weight of the  
275 amino acids. For instance, the concentration of valine drops down to  $4.7 \times 10^{-8}$  M. These results corroborate  
276 previous amino acid analyses in such samples, where the abundances of amino acids in the residue decreases  
277 approximately with increasing carbon number <sup>36</sup>.

278 Compared with previous experiments, in 2002, from an ice composed of H<sub>2</sub>O:CH<sub>3</sub>OH:NH<sub>3</sub>:CO:CO<sub>2</sub> in ratio  
279 2:1:1:1:1, Muñoz-Caro *et al.* identified 16 amino acids with a chiral GC-MS analyses and a similar sample  
280 pretreatment <sup>36</sup>. Even with a different initial ice composition, ratios related to glycine, the most abundant amino  
281 acid detected in all analyses, present a similar trend (Table 6). Other analyses performed on such organic  
282 residues present the same amino acid evolution <sup>37,38</sup>, but with lower diversities, especially when liquid  
283 chromatography technics are used <sup>37,39,40</sup>. Only GC×GC-TOF-MS analyses provide the detection of a same  
284 amino acid diversity with amino acid identification ranging from 24 to 29 amino acids depending on the  
285 derivatization procedure used <sup>31,41</sup>. GC-FT-Orbitrap-MS for amino acid targeted analyses offer thus comparable  
286 identification and quantification results to previous works.

### 287 4.2 GC-FT-orbitrap MS amino acid analysis of pre-cometary analog compared to GC×GC-TOF-MS and 288 UPLC-FT-orbitrap

289 To compare the identification procedure between different analytical instruments, the same residue as the one  
290 analyzed with GC-FT-Orbitrap-MS was analyzed with GC×GC-TOF-MS and UPLC-FT-Orbitrap-MS. **It has to**  
291 **be noted that for the GC-FT-Orbitrap-MS and GC×GC-TOF-MS methods, the same derivatization**  
292 **procedure is used, while with UPLC-FT-Orbitrap-MS, the samples are injected without derivatization.** As  
293 displayed in Table 7, identifications with UPLC-FT-Orbitrap-MS is less straight-forward than with GC-MS.  
294 Whereas in GC-MS, whatever are the technique used, 24 to 32 amino acid enantiomers present in the database  
295 were identified unambiguously, with UPLC-FT-Orbitrap-MS only 17 unambiguous identifications are obtained.  
296 These differences are mainly due to the poor resolution in LC compared to GC (Figure 4), the matrix effect that  
297 decreases the detectability of UPLC compared to GC that uses selective derivatization as well as the absence of  
298 reference standards for unambiguous identification. Since several amino acids present isomers that partially co-  
299 elute, their identification is problematic. Figure 4 displays the case of alanine isomers. No co-elution is observed

300 with GC-FT-orbitrap-MS, whereas L-alanine and  $\beta$ -alanine co-elute in the UPLC-FT-Orbitrap-MS analysis.  
301 While GC-FT-Orbitrap-MS in EI mode allowed for the unambiguous identification of L-alanine and  $\beta$ -alanine,  
302 UPLC-FT-Orbitrap-MS in ESI mode did not allow for their identification as they share the same molecular ion  
303 (Figure S1). However, MS/MS can also be performed with UPLC-FT-Orbitrap, thus providing structural  
304 information and allowing the discrimination of isomers. However, for this strategy to be efficient, higher amount  
305 of pre-cometary analog material should be available.

## 306 5. CONCLUSION

307 In this contribution, we investigated the interest of coupling gas chromatography to high-resolution mass  
308 spectrometry, a FT-Orbitrap-MS, presenting a resolution of 60 000 at  $m/z$  200. This system was tested for the  
309 characterization of derivatized amino acids in a laboratory produced sample analog to pre-cometary organic  
310 matter. Ionization modes were first tested and optimized. The molecular ion was only obtained with enough  
311 intensities in the positive chemical ionization mode with  $\text{NH}_3$  or  $\text{CH}_4/\text{NH}_3$  (90/10) reactant gas. Interestingly, the  
312 negative chemical ionization mode using  $\text{NH}_3$  as a reactant gas provided the molecular ion specifically of  $\alpha$ -  
313 amino acids. The electron impact mode remains the most sensitive ionization mode, offering LOD and LOQ  
314 values in the range of  $2 \times 10^{-7}$  M and  $6 \times 10^{-7}$  M, respectively. We observed a variability in the LOD values as  
315 a consequence of the difference in amino acid selectivity of the derivatization method used. GC-FT-Orbitrap-MS  
316 results were compared with two other analytical instruments: GC $\times$ GC-TOF-MS and UPLC-FT-Orbitrap-MS.  
317 Same compound identifications as GC $\times$ GC-TOF-MS are observed, while with UPLC-MS chromatographic  
318 resolution is weaker than with GC-MS, making amino acid identifications more challenging due to increased co-  
319 elution of amino acid isomers. Hence, identification of targeted amino acids using UPLC-MS without  
320 derivatization tend to be more limited than with GC-MS methodologies.

321 From all these considerations, it appears that the GC-FT-Orbitrap-MS is an attractive system for targeted  
322 analyses of amino acids in unknown samples. The EI mode directly provides information on the chemical  
323 structures, enhancing compound identification, while the optimized resolution enables the separation of  
324 enantiomers and isomers. Furthermore, the high resolution of the orbitrap allows to perform mass extraction in  
325 full scan that provides a  $S/N$  near zero, allowing high sensitivities. It also provides formula attribution to ions of  
326 fragment pattern with error lower than 2 ppm, giving crucial information for identification of compounds in  
327 samples. GC-FT-Orbitrap-MS is thus a powerful system for the targeted analysis of compounds **in samples**  
328 **presenting a high organic molecular diversity as demonstrate with the searched for amino acids in pre-**  
329 **cometary analogs.**

330

## 331 Acknowledgement

332 The research was funded by the Centre National d'Etudes Spatiales (CNES, R-S18/SU-0003-072 and R-S18/SU-  
333 0003-072, PI: G. D.), and the Centre National de la Recherche Française (CNRS) with the programs "Physique  
334 et Chimie du Milieu Interstellaire" (PCMI-PI:G.D.) and "Programme National de Planétologie" (PNP) (PI:  
335 G.D.). G.D is grateful to the Agence Nationale de la Recherche for funding via the ANR RAHIIA\_SSOM  
336 (ANR-16-CE29-0015) and VAHIIA (ANR-12-JS08-0001). The project has further received funding from the  
337 EXCellence Initiative of Aix-Marseille Université - A\*Midex, a French "Investissements d'Avenir programme"  
338 AMX-21-IET-018, from the Région SUD Provence Alpes Côte d'Azur "Apog 2017" – PILSE, as well by the EU  
339 Framework Program for Research and Innovation Horizon 2020 (grants ERC 804144).

340

## 341 Reference

- 342 1. Abreu, N. M., Aponte, J. C., Cloutis, E. A. & Nguyen, A. N. The Renazzo-like carbonaceous chondrites  
343 as resources to understand the origin, evolution, and exploration of the solar system. *Geochemistry* **80**,  
344 125631 (2020).
- 345 2. Glavin, D. P. *et al.* *The Origin and Evolution of Organic Matter in Carbonaceous Chondrites and Links*  
346 *to Their Parent Bodies. Primitive Meteorites and Asteroids* (Elsevier Inc., 2018). doi:10.1016/B978-0-  
347 12-813325-5.00003-3.
- 348 3. Caselli, P. & Ceccarelli, C. Our astrochemical heritage. *Astron. Astrophys. Rev.* **20**, 56 (2012).

- 349 4. Biver, N. *et al.* Molecular composition of short-period comets from millimetre-wave spectroscopy:  
350 21P/Giacobini-Zinner, 38P/Stephan-Oterma, 41P/Tuttle-Giacobini-Kresák, and 64P/Swift-Gehrels.  
351 *Astron. Astrophys.* **651**, A25 (2021).
- 352 5. Rubin, M., Bekaert, D., Broadley, M., Drozdovskaya, M. N. & Wampfler, S. Volatile species in comet  
353 investigating the link from the ISM to the terrestrial planets. *ACS Earth Sp. Chem.* (2019)  
354 doi:10.1021/acsearthspacechem.9b00096.
- 355 6. Biver, N. *et al.* Ethyl alcohol and sugar in comet C / 2014. *Sci. Adv.* (2015) doi:10.1126/sciadv.1500863.
- 356 7. Capaccioni, F. *et al.* The organic-rich surface of comet 67P/Churyumov-Gerasimenko as seen by  
357 VIRTIS/Rosetta. *Science (80-. )*. **347**, 0628 (2014).
- 358 8. Quirico, E. *et al.* Refractory and semi-volatile organics at the surface of comet 67P/Churyumov-  
359 Gerasimenko: Insights from the VIRTIS/Rosetta imaging spectrometer. *Icarus* **272**, 32–47 (2016).
- 360 9. Poch, O. *et al.* Ammonium salts are a reservoir of nitrogen on a cometary nucleus and possibly on some  
361 asteroids. *Science (80-. )*. **367**, eaaw7462 (2020).
- 362 10. Keller, H. U. & Kührt, E. Cometary Nuclei—From Giotto to Rosetta. *Space Sci. Rev.* **216**, 14 (2020).
- 363 11. Altwegg, K. *et al.* Prebiotic chemicals — amino acid and phosphorus — in the coma of comet 67P /  
364 Churyumov-Gerasimenko. *Science (80-. )*. **2**, 1–6 (2016).
- 365 12. Schmitt-Kopplin, P. *et al.* High molecular diversity of extraterrestrial organic matter in Murchison  
366 meteorite revealed 40 years after its fall. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2763–2768 (2010).
- 367 13. Elsila, J. E. *et al.* Meteoritic Amino Acids: Diversity in Compositions Reflects Parent Body Histories.  
368 *ACS Cent. Sci.* acscentsci.6b00074 (2016) doi:10.1021/acscentsci.6b00074.
- 369 14. Selliez, L. *et al.* High- resolution mass spectrometry for future space missions: Comparative analysis of  
370 complex organic matter with LAb- CosmOrbitrap and laser desorption/ionization Fourier transform ion  
371 cyclotron resonance. *Rapid Commun. Mass Spectrom.* **34**, (2020).
- 372 15. Briois, C. *et al.* Orbitrap mass analyser for in situ characterisation of planetary environments:  
373 Performance evaluation of a laboratory prototype. *Planet. Space Sci.* **131**, 33–45 (2016).
- 374 16. Danger, G. *et al.* Unprecedented Molecular Diversity Revealed in Meteoritic Insoluble Organic Matter:  
375 The Paris Meteorite’s Case. *Planet. Sci. J.* **1**, 55 (2020).
- 376 17. Danger, G. *et al.* The transition from soluble to insoluble organic matter in interstellar ice analogs and  
377 meteorites. *Astron. Astrophys.* (2022) doi:10.1051/0004-6361/202244191.
- 378 18. Hashiguchi, M. & Naraoka, H. High-mass resolution molecular imaging of organic compounds on the  
379 surface of Murchison meteorite. *Meteorit. Planet. Sci.* **17**, 1–17 (2018).
- 380 19. Callahan, M. P., Martin, M. G., Burton, A. S., Glavin, D. P. & Dworkin, J. P. Amino acid analysis in  
381 micrograms of meteorite sample by nanoliquid chromatography-high-resolution mass spectrometry. *J.*  
382 *Chromatogr. A* **1332**, 30–34 (2014).
- 383 20. Eddhif, B. *et al.* Development of liquid chromatography high resolution mass spectrometry strategies for  
384 the screening of complex organic matter: Application to astrophysical simulated materials. *Talanta* **179**,  
385 238–245 (2018).
- 386 21. Ruf, A., Poinot, P., Geffroy, C., D’Hendecourt, L. L. S. & Danger, G. Data-driven uplc-orbitrap ms  
387 analysis in astrochemistry. *Life* **9**, 1–14 (2019).
- 388 22. Seaton, K. M., Cable, M. L. & Stockton, A. M. Analytical Chemistry in Astrobiology. *Anal. Chem.* **93**,  
389 5981–5997 (2021).
- 390 23. Mahaffy, P. R. *et al.* The Sample Analysis at Mars Investigation and Instrument Suite. *Space Sci. Rev.*  
391 **170**, 401–478 (2012).
- 392 24. Pietrogrande, M. C. & Basaglia, G. Enantiomeric resolution of biomarkers in space analysis: Chemical  
393 derivatization and signal processing for gas chromatography–mass spectrometry analysis of chiral amino  
394 acids. *J. Chromatogr. A* **1217**, 1126–1133 (2010).

- 395 25. Poinot, P. & Geffroy-Rodier, C. Searching for organic compounds in the Universe. *TrAC Trends Anal.*  
396 *Chem.* **65**, 1–12 (2015).
- 397 26. Arevalo, R., Ni, Z. & Danell, R. M. Mass spectrometry and planetary exploration: A brief review and  
398 future projection. *J. Mass Spectrom.* **55**, (2020).
- 399 27. Peterson, A. C. *et al.* Development of a GC/quadrupole-orbitrap mass spectrometer, Part I: Design and  
400 characterization. *Anal. Chem.* **86**, 10036–10043 (2014).
- 401 28. Peterson, A. C., Balloon, A. J., Westphall, M. S. & Coon, J. J. Development of a GC/quadrupole-  
402 orbitrap mass spectrometer, part II: New approaches for discovery metabolomics. *Anal. Chem.* **86**,  
403 10044–10051 (2014).
- 404 29. Javelle, T., Riguezza, M. & Danger, G. Identify low mass volatile organic compounds from cometary ice  
405 analogs using gas chromatography coupled to an Orbitrap mass spectrometer associated to electron and  
406 chemical ionizations. *J. Chromatogr. A* **1652**, 462343 (2021).
- 407 30. Gudipati, M. S. *et al.* Laboratory Studies Towards Understanding Comets. *Space Sci. Rev.* **197**, 101–150  
408 (2015).
- 409 31. Meinert, C. *et al.* N-(2-Aminoethyl)glycine and Amino Acids from Interstellar Ice Analogues. *Chem.*  
410 *Plus Chem.* **77**, 186–191 (2012).
- 411 32. Myrgorodska, I., Meinert, C., Martins, Z., le Sergeant d’Hendecourt, L. & Meierhenrich, U. J.  
412 Quantitative enantioseparation of amino acids by comprehensive two-dimensional gas chromatography  
413 applied to non-terrestrial samples. *J. Chromatogr. A* **1433**, 131–136 (2016).
- 414 33. Meinert, C. & Meierhenrich, U. J. Derivatization and Multidimensional Gas-Chromatographic  
415 Resolution of  $\alpha$ -Alkyl and  $\alpha$ -Dialkyl Amino Acid Enantiomers. *Chempluschem* **79**, 781–785 (2014).
- 416 34. Serra, C. *et al.* Integrative analytical workflow to enhance comprehensive analysis of organic molecules  
417 in extraterrestrial objects. *Talanta* **243**, 123324 (2022).
- 418 35. Freissinet, C. *et al.* Search for evidence of life in space: Analysis of enantiomeric organic molecules by  
419 N,N-dimethylformamide dimethylacetal derivative dependant Gas Chromatography–Mass Spectrometry.  
420 *J. Chromatogr. A* **1217**, 731–740 (2010).
- 421 36. Muñoz Caro, G. M. *et al.* Amino acids from ultraviolet irradiation of interstellar ice analogues. *Nature*  
422 **416**, 403–406 (2002).
- 423 37. Nuevo, M., Auger, G., Blanot, D. & d’Hendecourt, L. A detailed study of the amino acids produced from  
424 the vacuum UV irradiation of interstellar ice analogs. *Orig. Life Evol. Biosph.* **38**, 37–56 (2008).
- 425 38. Nuevo, M. *et al.* The effects of circularly polarized light on amino acid enantiomers produced by the UV  
426 irradiation of interstellar ice analogs. *Astron. Astrophys.* **457**, 741–751 (2006).
- 427 39. Bernstein, M. P., Dworkin, J. P., Sandford, S. A., Cooper, G. W. & Allamandola, L. J. Racemic amino  
428 acids from the ultraviolet photolysis of interstellar ice analogues. *Nature* **416**, 401–403 (2002).
- 429 40. Elsila, J. E., Dworkin, J. P., Bernstein, M. P., Martin, M. P. & Sandford, S. A. Mechanisms of amino  
430 acid formation in interstellar ice analogs. *Astrophys. J.* **660**, 911–918 (2007).
- 431 41. Modica, P. *et al.* Enantiomeric Excesses Induced in Amino Acids By Ultraviolet Circularly Polarized  
432 Light Irradiation of Extraterrestrial Ice Analogs: a Possible Source of Asymmetry for Prebiotic  
433 Chemistry. *Astrophys. J.* **788**, 79 (2014).

434

## 435 FIGURES AND TABLES

436 **Table 1 - Amino acids of the standard solution.**

Amino acid group	Amino acids
N-alkyl	sarcosine, N-ethylglycine
$\alpha$ -amino acids	DL-2-ABA, DL-alanine, DL- <i>allo</i> -isoleucine, DL-asparagine, DL-aspartic acid, DL-cysteine, DL-glutamic acid, glycine, DL-isoleucine, DL-leucine, DL-

methionine, DL-norleucine, DL-norvaline, DL-phenylalanine, DL-proline, DL-serine, DL-valine

$\alpha$ ,  $\alpha$ -dialkyl amino acids 2-AIBA

$\beta$ -amino acids  $\beta$ -alanine, DL-3-AIBA, DL-3-ABA, DL- $\beta$ -leucine

$\gamma$ -amino acids  $\gamma$ -aminobutyric acid (GABA)

$\delta$ -amino acid 5-aminovaleic acid (5-AVA)

diamino acids DL-2,3-DAPA, DL-2,4-DABA

437

438 **Table 2** - Average retention time of 47 amino acids, theoretical  $m/z$  ratios of the base peak and their average  
 439 error at the theoretical value in ppm. Are also reported the stoichiometric formula of the base peak and the part  
 440 of the derivatization agent is marked in orange.

N°	Amino acids ECHFBE	Retention time <sup>b</sup>		Base peak formula	Theoretical $m/z$	$m/z$ error <sup>c</sup> (ppm) +-
		$t_R$ (min)	$\sigma$			
1	sarcosine	22.20	0.22	$C_2H_5N_1C_3H_5O_2$	116.07061	1.10
2	<i>N</i> -ethylglycine	25.54	0.16	$C_3H_7N_1C_3H_5O_2$	130.08626	1.12
3	2-AIBA	30.30	0.02	$C_3H_7N_1C_3H_5O_2$	130.08626	1.38
4	D-alanine	32.08	0.08	$C_2H_5N_1C_3H_5O_2$	116.07061	0.95
5	L-alanine	33.27	0.07			
6	glycine	34.67	0.17	$C_1H_3N_1C_3H_5O_2$	102.05496	0.88
7	D-2-ABA	35.71	0.08	$C_3H_7N_1C_3H_5O_2$	130.08626	0.85
10	L-2-ABA	36.82	0.07			
9	proline	36.47	0.18	$C_4H_7N_1C_3H_5O_2$	142.08626	0.70
8	D-valine	36	0.07	$C_4H_9N_1C_3H_5O_2$	144.10191	0.82
11	L-valine	36.93	0.06			
12	$\beta$ -alanine	37.28	0.14	$C_1H_3N_1C_3H_5O_2$	102.05496	0.97
13	3-AIBA	37.87	0.12	$C_1H_3N_1C_3H_5O_2$	102.05496	1.03
14	D-3-ABA	39.15	0.06	$C_2H_5N_1C_3H_5O_2$	116.07061	0.99
15	L-3-ABA	39.43	0.06			
16		39.77	0.07			
18	isoleucine isomers <sup>a</sup>	40.70	0.06	$C_5H_{11}N_1C_3H_5O_2$	158.11756	1.02
19		41.02	0.07			
21		41.77	0.07			
17	D-norvaline	40.52	0.08	$C_4H_9N_1C_3H_5O_2$	144.10191	1.01
20	L-norvaline	41.61	0.07			
22	D-leucine	43.08	0.08	$C_1H_3N_1C_3H_5O_2$	102.05496	1.09
23	L-leucine	44.20	0.07			
24	D-norleucine	45.15	0.11	$C_5H_{11}N_1C_3H_5O_2$	158.11756	0.94
26	L-norleucine	46.11	0.09			
25	D- $\beta$ -leucine	45.48	0.12	$C_2H_4N_1C_3H_5O_2C_5H_2F_7O_2$	342.05708	0.93
27	L- $\beta$ -leucine	46.11	0.11			
28	GABA	46.42	0.06	$C_4H_9C_3H_2F_7O_2$	284.05184	0.98
29	D-aspartic acid	50.58	0.14	$C_3H_4N_1C_7H_7O_4F_7$	342.05708	1.14
30	L-aspartic acid	50.95	0.13			
31	5-AVA	52.06	0.01	$C_4H_9N_1C_5H_2O_2F_7$	298.06725	0.48
32	D-serine	56.14	0.15	$C_1H_2NC_3H_5O_2$	114.05485	1.05
33	L-serine	56.57	0.14			
34	D-asparagine	57.44	0.18	$C_3H_4N_2C_3H_5O_2$	141.06585	0.94
35	L-asparagine	57.66	0.01			
36	D-methionine	58.41	0.45	$C_1H_2NC_3H_5O_2$	101.04713	1.03
38	L-methionine	59.17	0.39			
37	D-glutamic acid	58.41	0.20	$C_1H_1NC_3H_5O$	84.04459	0.68
39	L-glutamic acid	59.17	0.16			
40	D-phenylalanine	61.04	0.20	$C_7H_7$	91.05423	1.01
41	L-phenylalanine	61.55	0.17			
42	D-cysteine	63.64	0.04	$C_3H_4ONS$	102.00081	1.04
43	L-cysteine	63.99	0.05			

44	D-2,3-DAPA	72.37	0.19	CH <sub>3</sub> NC <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	102.05496	1.10
45	L-2,3-DAPA	73.28	0.17			
46	D-2,4-DABA	76.52	0.02	C <sub>3</sub> H <sub>5</sub> NC <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	128.07061	0.84
47	L-2,4-DABA	76.87	0.01			

441 <sup>a</sup> Due to the lack of individual standards the four stereoisomers of isoleucine (DL-isoleucine and DL-*allo*-  
442 isoleucine) were not identified.

443 <sup>b</sup> Averages on retention times are obtained from 11 values.

444 <sup>c</sup> m/z error = (measured-theory/theory)x10<sup>6</sup>.

445 **Table 3** - Performance of different chemical ionization modes and gases used. The percentage of the molecular  
446 ion derived from the abundance of the molecular ion in one mode divided by the sum of the abundances of all  
447 studied ionization modes for the same amino acid. Only D-amino acids are represented for simplicity, except for  
448 the four forms of isoleucine (DL-isoleucine and DL-*allo*-isoleucine). Amino acids with a molecular ion mass  
449 ((M+H)<sup>+</sup>) greater than 400 m/z could not be detected due to the mass range of the method of 50-400 amu.

Amino acids ECHFBE	NCI			PCI			EI	
	CH <sub>4</sub> relative %	NH <sub>3</sub> relative %	NH <sub>3</sub> /CH <sub>4</sub> relative %	CH <sub>4</sub> relative %	NH <sub>3</sub> relative %	NH <sub>3</sub> /CH <sub>4</sub> relative %	(70eV) relative %	
sarcosine	0.01	8.06	0.44	2.59	16.61	51.35	20.93	
<i>N</i> -ethylglycine	0.02	12.57	0.47	2.64	25.59	49.61	9.11	
alanine	0.40	43.50	0.75	3.16	13.20	39.00	-	
glycine	0.23	14.89	0.44	5.55	10.38	42.14	26.36	
valine	0.03	50.50	1.54	1.01	13.21	33.71	-	
proline	0.00	3.33	0.15	2.57	30.74	59.42	3.78	
leucine	0.39	29.44	1.30	2.10	17.01	48.58	1.18	
isoleucine	1	0.31	39.50	1.34	0.98	16.48	41.40	-
	2	0.33	38.96	1.51	1.13	17.07	41.01	-
	3	0.40	44.45	1.55	0.91	15.35	37.34	-
	4	0.39	44.23	1.70	1.03	15.57	37.08	-
cysteine	0.95	5.91	1.99	2.13	5.91	2.59	80.52	
2-ABA	0.43	52.09	1.19	1.86	11.01	33.42	-	
norvaline	0.44	40.33	1.28	2.04	15.45	40.07	0.39	
norleucine	0.50	33.14	1.57	2.07	15.00	46.34	1.39	
β-alanine	-	0.16	-	1.80	19.49	37.26	41.29	
β-leucine	-	-	-	1.71	33.71	64.15	0.42	
3-ABA	-	-	-	1.61	28.59	55.42	14.38	
3-AIBA	-	0.98	-	1.51	25.46	52.93	19.12	
GABA	-	-	-	0.00	24.28	61.56	14.16	

450 - : not identified

451

452 **Table 4** – LOD and LOQ determined for each amino acid, ranked from lowest to highest detection threshold.

Amino acids ECHFBE		LOQ (×10 <sup>-7</sup> M)	LOD (×10 <sup>-7</sup> M)
sarcosine		4.99	1.50
isoleucine & <i>allo</i> - isoleucine isomers	1	5.04	1.66
	2	5.70	1.88
	3	5.42	1.79
	4	7.04	2.32
norleucine	D	5.19	1.71
	L	6.45	2.13
valine	D	5.91	1.95
	L	6.83	2.26
norvaline	D	6.19	2.04
	L	6.70	2.21
2-ABA	D	6.35	2.10

	L	6.76	2.23
leucine	D	6.76	2.23
	L	6.96	2.30
DL-proline		6.97	2.30
phenylalanine	D	7.99	2.64
	L	11.2	3.69
alanine	D	9.47	3.12
	L	9.91	3.27
N-ethylglycine		11.8	3.55
$\beta$ -alanine		12.6	4.02
3-AIBA		22.4	6.71
aspartic acid	D	31.3	10.3
	L	25.1	8.28
glycine		26.7	8.82
$\beta$ -leucine	D	39.0	12.9
	L	39.2	12.9
2,3-DAPA	D	42.3	14.0
	L	40.2	13.3
serine	D	64.4	21.3
	L	58.8	19.4
methionine	D	70.1	23.1
	L	60.6	20.0
3-ABA	D	79.9	26.4
	L	75.1	24.8

453

454 **Table 5** – Comparison of amino acid response obtained from SIM to full scan mode for the 43 amino acid  
455 standards at a concentration of  $10^{-5}$  M. Mass extraction is used with SIM when the mass windows overlap,  
456 which prevents an identification of the amino acid with the corresponding characteristic mass.

Amino acids	<i>m/z</i> of ions used for SIM or mass extraction	$A_x/A_{IS}$			EM/Full scan vs EM/SIM			
		Full scan with mass extraction ( $\times 10^{-1}$ )	SIM ( $\times 10^{-1}$ )	SIM with mass extraction ( $\times 10^{-1}$ )	$A_x/A_{IS}$ average ( $\times 10^{-1}$ )	$\sigma$ ( $\times 10^{-1}$ )	% $\sigma$	
sarcosine	116.07061	12.2	11.9	13.1	12.6	0.45	4	
N-ethylglycine	130.08626	5.99	7.40	8.29	7.14	1.15	16	
2-AIBA	130.08626	0.19	0.38	0.453	0.32	0.13	41	
alanine	116.07061	D	7.72	8.23	8.77	8.24	0.52	6
		L	8.10	8.83	9.13	8.61	0.52	6
glycine	102.05496	4.42	6.67	7.00	5.71	1.29	23	
2-ABA	130.08626	D	9.50 #		7.06	8.28	1.22	15
		L	11.8 #		5.83	8.81	2.98	34
DL-proline	142.08626	40.5 #		17.2	28.8	11.7	40	
valine	144.10191	D	11.2 #		5.76	8.46	2.70	32
		L	13.0 #		4.53	8.78	4.25	48
$\beta$ -alanine	98.02358	1.18 #		0.83	1.00	0.18	18	
3-AIBA	102.05496	4.49	3.62	3.79	4.14	0.35	8	
3-ABA	116.07061	D	0.64 #		0.29	0.46	0.18	38
		L	0.52 #		0.21	0.37	0.16	43
isoleucine	158.11756	1	3.34	1.25	1.24	2.29	1.05	46
		2	3.54	1.26	1.33	2.44	1.10	45
		3	5.44	2.00	2.05	3.74	1.70	45
		4	6.01	2.41	2.30	4.16	1.86	45
norvaline	144.10191	D	11.2	5.73	6.00	8.60	2.60	30
		L	11.7	6.00	6.39	9.06	2.67	29
leucine	158.11756	D	8.90	6.31	6.71	7.81	1.09	14

		L	9.02	6.69	6.96	7.99	1.03	13
norleucine	158.11756	D	11.6	8.56	8.90	10.2	1.33	13
		L	13.0	3.22	3.28	8.16	4.88	60
β-leucine	342.05708	D	1.41 #		0.54	0.98	0.43	44
		L	1.35 #		0.49	0.92	0.43	47
GABA	284.05184		0.059	0.093	0.11	0.082	0.023	28
aspartic acid	254.00439	D	5.20	7.09	7.09	6.14	0.94	15
		L	5.28	6.94	7.46	6.37	1.09	17
5-AVA	298.06725		0.039	0.057	0.064	0.052	0.012	23
serine	114.05496	D	0.078	0.084	0.087	0.082	0.005	6
		L	0.083	0.081	0.091	0.087	0.004	5
asparagine	69.04477	D	2.15 #		0.63	1.39	0.76	55
		L	2.67 #		0.75	1.71	0.96	56
phenylalanine	91.05423	D	18.9	34.7	35.2	27.0	8.14	30
		L	22.1	40.3	42.0	32.1	9.93	31
cysteine	85.98207	D	2.38	3.58	2.99	2.69	0.31	11
		L	2.47	3.80	3.20	2.83	0.37	13
2,3-DAPA	102.05496	D	6.58	6.67	6.72	6.65	0.067	1
		L	6.41	6.51	6.60	6.50	0.096	1
2,4-DABA	128.07061	D	0.99	0.75	0.76	0.88	0.11	13
		L	1.01	0.75	0.78	0.90	0.12	13

457 # amino acids that cannot be integrated due to poorly resolved or coeluted amino acid masses.

458

459 **Table 6** – Amino acids identified in the residue with the theoretical mass of the <sup>13</sup>C-labeled ECHFBE amino acid  
460 base peaks. In orange are noted the atoms corresponding to the derivatization agent. Also displayed are the  
461 concentrations of some identified amino acids. Number of mass fragments used for amino acid identification  
462 Table S3. n.q: not quantifiable.

N°	Amino Acids	$t_R$ (min)	Amino acid stoichiometric formula	Base peak ECHFBF amino acid	Theoretical mass of base peak		Concentration (M)	% relative to glycine	% relative to glycine		
					$^{12}C$	$^{13}C$		This work	Muñoz- Caro 2002		
1	sarcosine	22.15	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	C <sub>2</sub> H <sub>5</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	116.07061	118.07732	2.09×10 <sup>-5</sup>	3.9	5.71		
2	N-ethylglycine	25.52	C <sub>4</sub> H <sub>9</sub> NO <sub>2</sub>	C <sub>3</sub> H <sub>7</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	130.08626	133.09632	1.12×10 <sup>-6</sup>	0.21	1.91		
3	alanine	D	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	C <sub>2</sub> H <sub>5</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	116.07061	118.07732	9.27×10 <sup>-5</sup>	17.29	19.3		
4		L								33.23	9.33×10 <sup>-5</sup>
5	glycine	34.65	C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub>	C <sub>1</sub> H <sub>3</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	102.05496	103.05831	5.36×10 <sup>-4</sup>	100	100		
6	2-ABA	D	C <sub>4</sub> H <sub>9</sub> NO <sub>2</sub>	C <sub>3</sub> H <sub>7</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	130.08626	133.09632	2.41×10 <sup>-6</sup>	0.45			
7		L								36.79	2.07×10 <sup>-6</sup>
8	DL-proline	36.5	C <sub>5</sub> H <sub>9</sub> NO <sub>2</sub>	C <sub>4</sub> H <sub>7</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	142.08626	146.09968	2.74×10 <sup>-7</sup>	0.05	0.06		
9	valine	D	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub>	C <sub>4</sub> H <sub>9</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	144.10191	148.11533	nq		0.61		
10		L								36.9	nq
11	β-alanine	37.27	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	C <sub>1</sub> H <sub>3</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	102.05496	103.05831	2.05×10 <sup>-5</sup>	3.82	4.29		
12	3-AIBA	37.87	C <sub>4</sub> H <sub>9</sub> NO <sub>2</sub>	C <sub>1</sub> H <sub>3</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	102.05496	103.05831	nq				
13	3-ABA	D	C <sub>4</sub> H <sub>9</sub> NO <sub>2</sub>	C <sub>2</sub> H <sub>5</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	116.07061	118.07732	1.67×10 <sup>-5</sup>	3.12			
14		L								39.43	1.62×10 <sup>-5</sup>
15	isoleucine	1	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	C <sub>5</sub> H <sub>11</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	158.11756	163.13433	nq				
16		2								40.71	nq
17		3								41.04	nq
18		4								41.78	nq
19	norvaline	D	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub>	C <sub>4</sub> H <sub>9</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	144.10191	148.11533	nq				
20		L								41.58	nq
21	leucine	D	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	C <sub>1</sub> H <sub>3</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	102.05496	103.05831	nq				
22		L								44.17	nq
23	norleucine	D	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	C <sub>5</sub> H <sub>11</sub> N <sub>1</sub> C <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	158.11756	163.13433	nq				
24		L								46.05	nq
25	aspartic acid	D	C <sub>4</sub> H <sub>7</sub> NO <sub>4</sub>	C <sub>3</sub> H <sub>4</sub> N <sub>1</sub> C <sub>7</sub> H <sub>7</sub> O <sub>4</sub> F <sub>7</sub>	342.05708	345.06715	nq		1.4		
26		L								50.89	nq
27	serine	D	C <sub>3</sub> H <sub>7</sub> NO <sub>3</sub>	C <sub>2</sub> H <sub>4</sub> N <sub>1</sub> C <sub>3</sub> H <sub>4</sub> O <sub>2</sub>	114.05496	116.06167	1.10×10 <sup>-5</sup>	2.05	3.29		
28		L								56.56	1.49×10 <sup>-5</sup>
29	2,3- DAPA	D	C <sub>3</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub>	CH <sub>3</sub> NC <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	102.05496	103.05832	6.43×10 <sup>-6</sup>	1.20	20		
30		L								73.2	6.48×10 <sup>-6</sup>
31	2,4- DABA	D	C <sub>4</sub> H <sub>10</sub> N <sub>2</sub> O <sub>2</sub>	C <sub>3</sub> H <sub>5</sub> NC <sub>3</sub> H <sub>5</sub> O <sub>2</sub>	128.07061	131.08068	nq				
32		L								76.79	nq

463

464 **Table 7** - Targeted amino acid identifications in a pre-cometary organic analog H<sub>2</sub>O:<sup>13</sup>CH<sub>3</sub>OH:NH<sub>3</sub> (3:1:1) using  
465 the three analytical techniques.

Amino acids	GC (ECHFBF)		GC×GC (ECHFBF)	UPLC
		FT-Orbitrap-MS	TOF-MS	FT-Orbitrap-MS
sarcosine	1	id	id	id
N-ethylglycine	2	id	id	id
2-AIBA		nd	#	und
D-alanine	3	id	id	id
L-alanine	4	id	id	und
Glycine	5	id	id	id
D-2-ABA	6	id	id	und
L-2-ABA	7	id	id	und

DL-proline	8	id	id	id
D-valine	9	id	id	und
L-valine	10	id	id	und
$\beta$ -alanine	11	id	id	und
3-AIBA	12	id	id	und
D-3-ABA	13	id	id	und
L-3-ABA	14	id	id	und
D-isoleucine	15	id	nd	a
L-isoleucine	16	id	nd	a
D- <i>allo</i> -isoleucine	17	id	nd	a
L- <i>allo</i> -isoleucine	18	id	nd	a
D-norvaline	19	id	id	und
L-norvaline	20	id	id	und
D-leucine	21	id	nd	a
L-leucine	22	id	nd	a
D-norleucine	23	id	nd	a
L-norleucine	24	id	nd	a
$\beta$ -leucine		nd	#	a
GABA		nd	#	und
D-aspartic acid	25	id	id	id
L-aspartic acid	26	id	id	id
5-AVA		nd	id	und
DL-serine	27 -	id	id	id
	28			
D-asparagine		nd	nd	id
L-asparagine		nd	nd	id
DL-glutamic acid		nd	id	und
phenylalanine		nd	nd	und
D-2,3-DAPA	29	id	id	a
L-2,3-DAPA	30	id	id	a
D-2,4-DABA	31	id	id	a
L-2,4-DABA	32	id	id	a

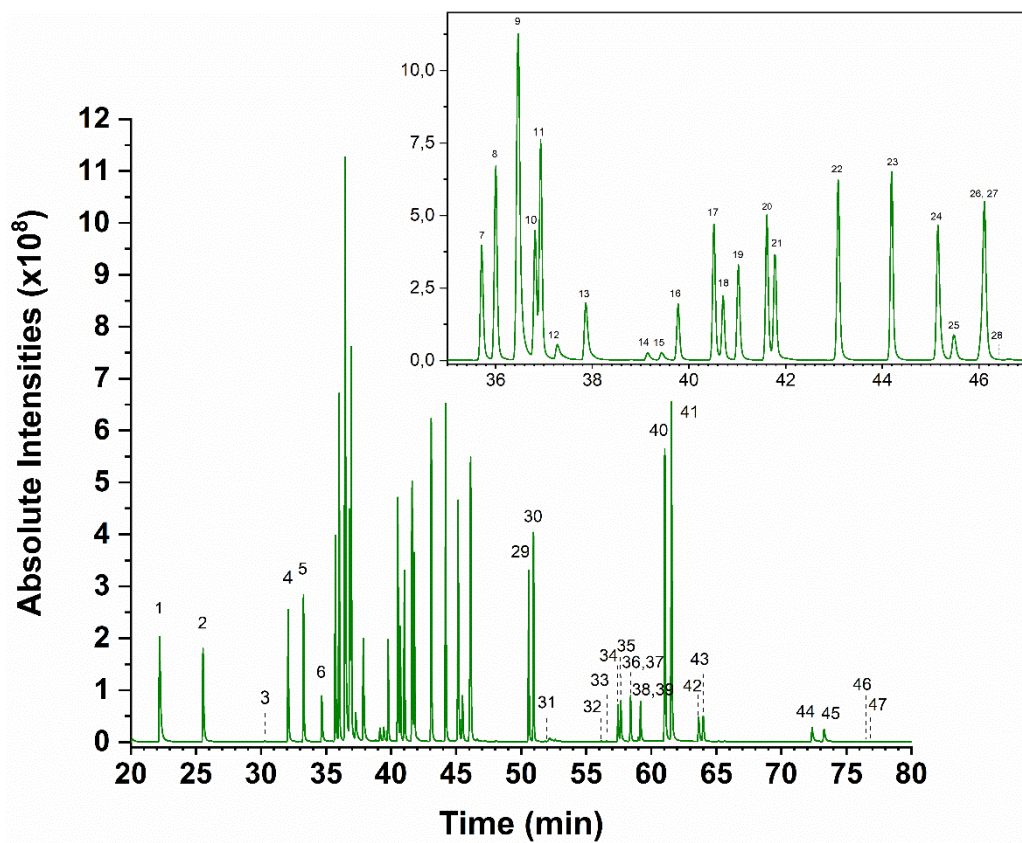
466 # tentatively identified only due to low *S/N* and consequently poor spectral match. <sup>a</sup> no standard available for  
467 confirmation. id: identified in the database. nd: not detected. und: undetermined due to coelution.

468

469

470

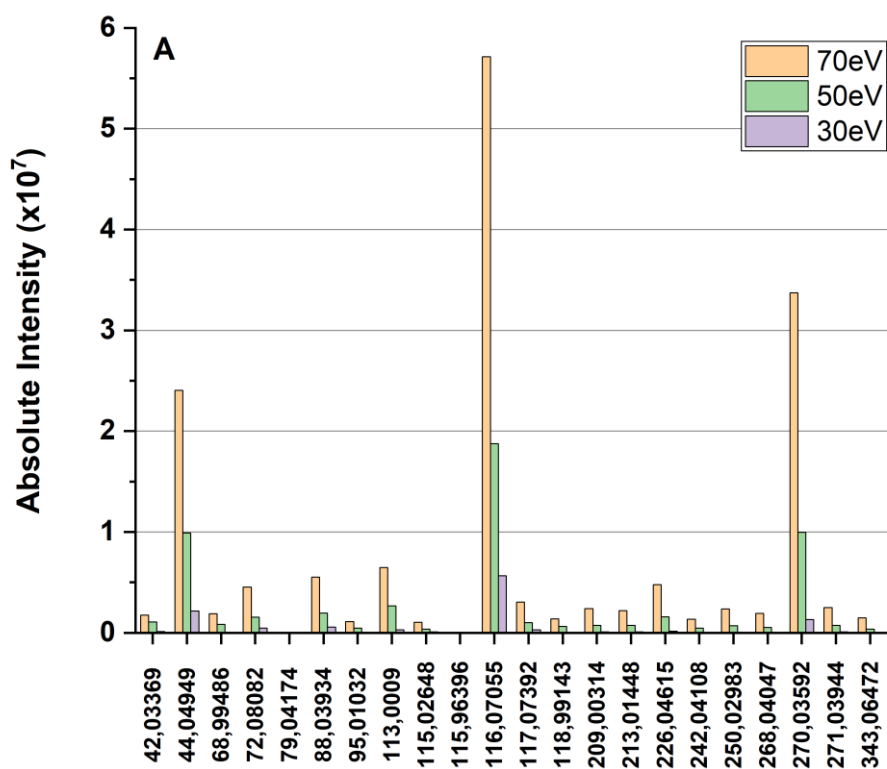
471



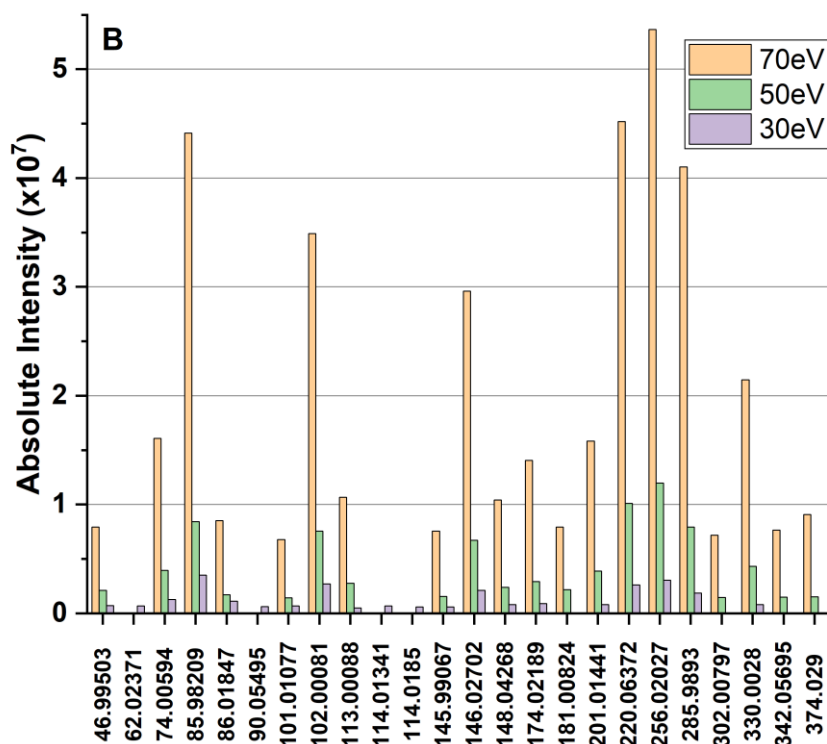
472

473 **Figure 1** – Reconstructed chromatogram obtained after mass extraction of the base peaks of each individual  
 474 amino acid (Table 2) present in the standard mixture at  $5 \times 10^{-5}$  M. Amino acid identification is given in Table  
 475 2. Analysis performed with EI mode at 70 eV.

476

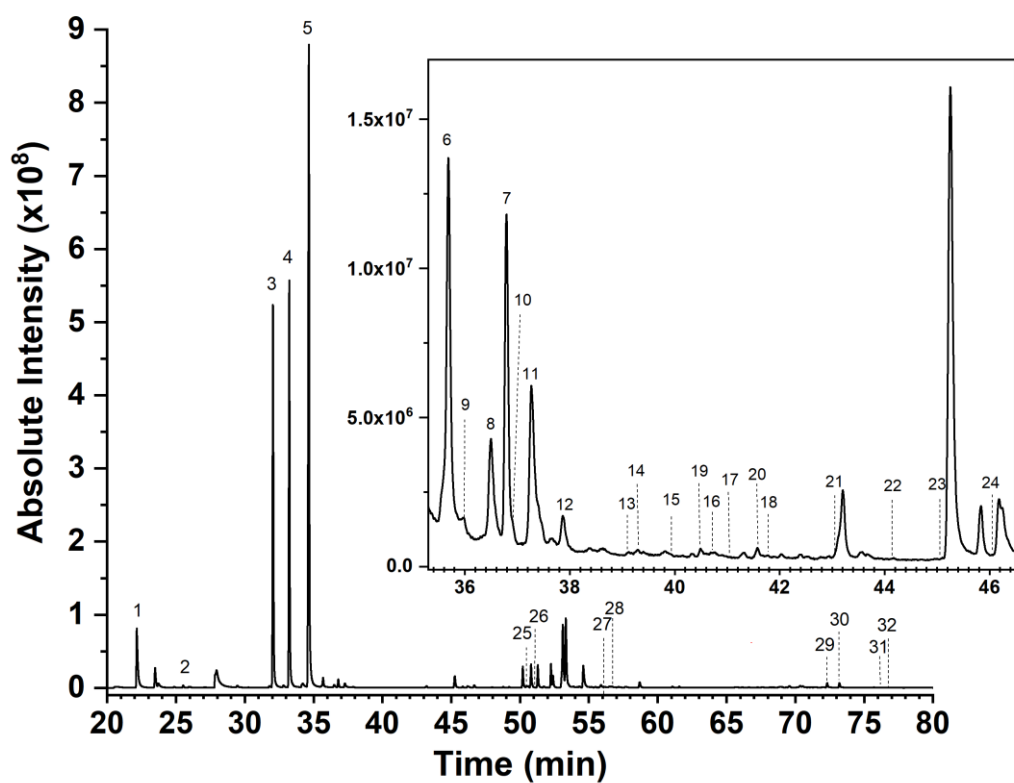


477



478

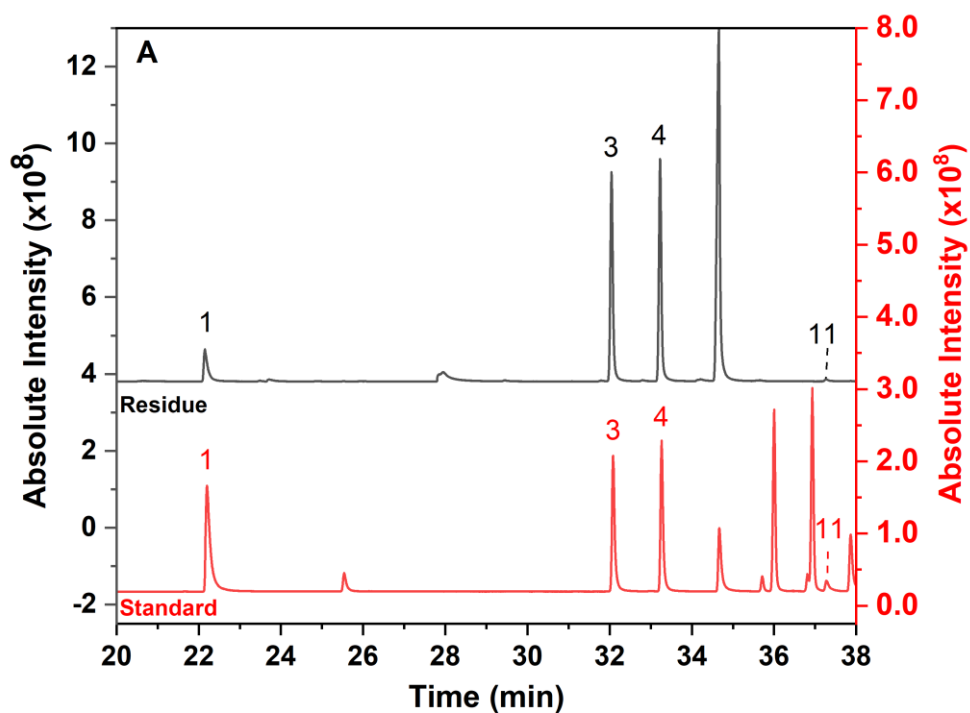
479 **Figure 2** - Evolution of the intensity of the observed fragment ions as a function of the ionization energy, (A) for  
 480 sarcosine (molecular ion at 343.06491 au, base peak at 116.07035 au), which is representative of the majority of  
 481 amino acids, and (B) for cysteine (molecular ion at 374.02915 au, base peak at 256.02027 au), which has a  
 482 particular fragmentation pattern compared to the other amino acids.



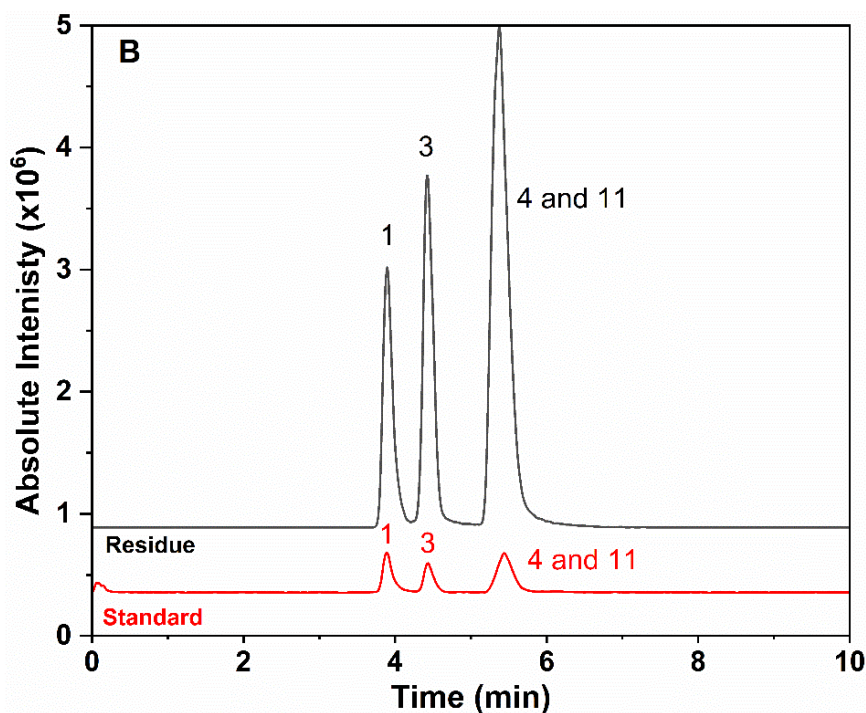
484

485 **Figure 3** – Reconstructed chromatogram obtained after mass extraction of the base peaks of each individual  
486 amino acid (Table S3). Amino acids identified in a pre-cometary analog using GC-FT-Orbitrap-MS after 24 h  
487 acid hydrolysis and subsequent derivatization. The derivatization procedure is presented in the Material and  
488 Method section.

489



490



491

492 **Figure 4** – Chromatograms displaying the analyses of alanine enantiomers (D-alanine 3 and L-alanine 4) and  
 493 alanine isomers, sarcosine (1) and  $\beta$ -alanine (11), in the residue and the reference standards with **A) GC-FT-**  
 494 **Orbitrap-MS after derivatization procedure and B) UPLC-FT-Orbitrap-MS without previous**  
 495 **derivatization.** The MS spectra of compounds L-alanine (4) and  $\beta$ -alanine (11) are reported in Figure S1 for the  
 496 UPLC-ESI-Orbitrap-MS and GC-EI-FT-Orbitrap-MS analyses.

497

498 amino acids, and (B) for cysteine (molecular ion at 374.02915 au, base peak at 256.02027 au), which has a  
 499 particular fragmentation pattern compared to the other amino acids.

500

501

502 **SUPPLEMENTARY INFORMATION**

503 **Table S1** - Retention time window and fragment ions (*m/z*) used in the SIM mode for each amino acid.

Compounds	Mass ( <i>m/z</i> )	Retention time window (min)
sarcosine	116.07061	22.5–26.0
<i>N</i> -ethylglycine	130.08626	26.0–32.0
2-AIBA	130.08626	26.0–32.0
alanine	116.07061	32.0–35.2
glycine	102.05496	35.2–36.3
2-ABA	130.08626	36.2–37.8
proline	142.08626	37.1–38.1
valine	144.10191	36.2–38.6
$\beta$ -alanine	98.02358	38.2–39.0
3-AIBA	102.05496	38.6–39.2
3-ABA	116.07061	39.5–41.0
isoleucine	158.11756	40.0–45.1
norvaline	144.10191	40.9–42.8
leucine	158.11756	40.0–45.1
norleucine	158.11756	45.6–47.2
$\beta$ -leucine	342.05708	45.8–47.2
GABA	284.05184	46.0–47.5
aspartic acid	254.00439	50.9–52.4
5-AVA	298.06725	52.7–53.5
serine	114.05496	56.6–57.7
asparagine	69.04477	57.5–59.5
methionine	101.04713	58.5–60.5
glutamic acid	84.04459	57.5–59.5
phenylalanine	91.05423	61.4–62.7
cysteine	85.98207	64.0–64.8
2,3-DAPA	102.05496	72.6–74.5
2,4-DABA	128.07061	76.2–79.0
IS methyl laurate	87.04401	39.5–40.3

504

505 **Table S2** – Concentration range used to estimate the LOD and LOQ of amino acids using GC-FT-Orbitrap-MS.

Concentration (M)			
$5 \times 10^{-7} - 2.5 \times 10^{-6}$	$7.5 \times 10^{-7} - 5 \times 10^{-6}$	$10^{-6} - 7.5 \times 10^{-6}$	$2 \times 10^{-6} - 10^{-5}$

sarcosine N-ethylglycine DL-alanine glycine DL-proline DL-valine DL-isoleucine DL- <i>allo</i> -isoleucine DL-leucine DL-phenylalanine DL-2-ABA DL-norvaline DL-norleucine	DL-3-AIBA	DL-aspartic acid $\beta$ -alanine DL- $\beta$ -leucine DL-3-ABA	DL-asparagine DL-methionine DL-serine DL-2.3-DAPA
--	-----------	--	--

506

507 **Table S3** – List of masses ( $m/z$ ) used for compound identification of amino acids in the organic residue of pre-  
508 cometary analog to pre-cometary matter. Number of mass fragments used for amino acid identification: a: 5, b:  
509 4, c: 3, d: 2, e: 1.

#	Amino Acids	RT (min)	Masse used for compound identification in residue vs standards	
			<sup>13</sup> C residue	<sup>12</sup> C standard
1	sarcosine <sup>a</sup>	22.15	118.07725 (100), 272.04258 (35.7), 90.04606 (9.62), 74.08753 (8.54), 228.05285 (4.93)	116.0705 (100), 270.03585 (33.5), 88.03924 (10.61), 72.08076 (9.85), 226.04592 (4.75)
2	N-ethylglycine <sup>a</sup>	25.52	133.09634 (98.58), 59.06853 (37.4), 287.06167 (24.73), 87.09977 (17.58), 103.05829 (12.84)	130.0862 (100), 58.06516 (38.08), 284.05137 (33.4), 86.09637 (18.79), 102.05488 (13.6)
3	D-alanine <sup>b</sup>	32.05	118.07719 (100), 272.04249 (30.19), 72.0355 (8.14), 74.08753 (7.79)	116.07055 (100), 270.03595 (30.15), 70.02879 (17.88), 72.08079 (8.58), 117.07388 (5.26)
4	L-alanine <sup>b</sup>	33.23		
5	glycine <sup>b</sup>	34.65	103.05822 (100), 257.02336 (77.36), 213.03368 (11.32), 57.01649 (7.02)	102.05491 (100), 256.0202 (75.64), 56.01312 (14.96), 212.03032 (9.92), 264.00902 (9.27)
6	D-2-ABA <sup>a</sup>	35.69	133.09636 (100), 61.07526 (35.61), 103.05832 (19.27), 87.05448 (15.64), 287.06172 (15.51)	130.08616 (100), 58.06515 (32.23), 284.05133 (18.09), 84.04437 (11.43), 102.0549 (7.08)
7	L-2-ABA <sup>a</sup>	36.79		
8	DL-proline <sup>c</sup>	36.5	146.09967 (3.26), 72.07187 (0.25), 116.06167 (0.14)	142.08615 (100), 70.0651 (36.39), 114.05493 (18.53), 296.05136 (15.38), 98.0964 (12.39)
9	D-valine <sup>c</sup>	35.97	148.11534 (42.53), 120.08408 (27.61), 103.01746 (21.99)	144.10175 (100), 116.0705 (71.08), 101.0107 (29.15), 98.06 (26.21), 256.0201 (15.78)
10	L-valine <sup>c</sup>	36.9		
11	$\beta$ -alanine <sup>a</sup>	37.27	273.04611 (10.53), 103.05829 (8.21), 117.06945 (8.21), 101.03375 (7.13), 72.03548 (4.79)	270.03559 (100), 115.06269 (68.9), 102.05486 (65.84), 98.02357 (57.56), 70.02875 (49.51)
12	3-AIBA <sup>a</sup>	37.87	103.0583 (100), 87.05446 (58.7), 116.05272 (36.36), 162.09454 (21.59), 132.08858 (20.84)	102.05487 (100), 84.04435 (74.81), 112.03922 (55.09), 158.08103 (33.04), 130.04984 (29.1)
13	D-3-ABA <sup>c</sup>	39.15	118.07736 (47.05), 288.06496 (20.26), 72.0355 (15.93)	116.07052 (100), 284.05138 (73.05), 227.02995 (36.09), 70.02878 (34.88), 270.03583 (30.51)
14	L-3-ABA <sup>c</sup>	39.43		
15	isoleucine forme 1 <sup>d</sup>	39.76	163.13424 (1.86), 103.0175 (1.66)	158.11741 (100), 101.0107 (61.1), 102.05489 (44.75), 283.00722 (28.36),

16	isoleucine forme 2 <sup>d</sup>	40.71		130.08623 (23.04)
17	isoleucine forme 3 <sup>d</sup>	41.04		
18	isoleucine forme 4 <sup>d</sup>	41.78		
19	D-norvaline <sup>b</sup>	40.49	148.11529 (40.01), 74.08755 (8.48), 258.02713 (5.05), 120.0841 (4.13)	144.10174 (100), 72.08076 (25.05), 256.02012 (13.33), 100.11199 (10.86), 116.07053 (8.62)
20	L-norvaline <sup>b</sup>	41.58		
21	D-leucine <sup>d</sup>	43.11	103.05821 (3.74), 163.13434 (1.84)	158.11737 (100), 102.05486 (92.64), 112.07561 (19.77), 256.02014 (17.91), 101.01071 (12.18)
22	L-leucine <sup>d</sup>	44.17		
23	D-norleucine <sup>e</sup>	45.1	163.13427 (9.04)	158.11737 (100), 86.09637 (28.4), 112.07562 (17.19), 102.05489 (14.69), 114.12767 (12.4)
24	L-norleucine <sup>e</sup>	46.05		
25	D-aspartic acid <sup>c</sup>	50.52	255.00804 (100), 345.06689 (69.69), 273.04604 (41.25)	254.00453 (100), 342.05658 (60.58), 270.03593 (35.4), 296.01504 (20.64), 228.02527 (17.84)
26	L-aspartic acid <sup>c</sup>	50.89		
27	D-serine <sup>b</sup>	56.14	116.0616 (100), 88.03037 (73.53), 103.04934 (29.3), 258.02706 (27.82)	86.0236 (100), 114.05491 (97.06), 256.02023 (27.2), 101.0107 (25.44), 314.02556 (24.72)
28	L-serine <sup>b</sup>	56.56		
29	D-2,3-DAPA <sup>b</sup>	72.31	103.05833 (100), 87.04643 (75.88), 159.06752 (45.38), 316.05186 (34.79)	102.05489 (100), 85.03961 (79.61), 157.06066 (44.89), 313.04148 (35.27), 101.01069 (28.47)
30	L-2,3-DAPA <sup>b</sup>	73.2		
31	D-2,4-DABA <sup>d</sup>	76.22	118.07713 (9.88), 131.08076 (7.28)	128.07066 (100), 116.07064 (64.37), 101.01078 (45.24), 129.04208 (28.62), 99.05531 (28.36)
32	L-2,3-DABA <sup>d</sup>	76.79		

510

511

#	Amino Acids	RT (min)	Masse used for compound identification in residue vs standards	
			<sup>13</sup> C residue	<sup>12</sup> C standard
1	sarcosine <sup>a</sup>	22.15	118.07725 (100), 272.04258 (35.7), 90.04606 (9.62), 74.08753 (8.54), 228.05285 (4.93)	116.0705 (100), 270.03585 (33.5), 88.03924 (10.61), 72.08076 (9.85), 226.04592 (4.75)
2	N-ethylglycine <sup>a</sup>	25.52	133.09634 (98.58), 59.06853 (37.4), 287.06167 (24.73), 87.09977 (17.58), 103.05829 (12.84)	130.0862 (100), 58.06516 (38.08), 284.05137 (33.4), 86.09637 (18.79), 102.05488 (13.6)
3	D-alanine <sup>b</sup>	32.05	118.07719 (100), 272.04249 (30.19), 72.0355 (8.14), 74.08753 (7.79)	116.07055 (100), 270.03595 (30.15), 70.02879 (17.88), 72.08079 (8.58), 117.07388 (5.26)
4	L-alanine <sup>b</sup>	33.23		
5	glycine <sup>b</sup>	34.65	103.05822 (100), 257.02336 (77.36), 213.03368 (11.32), 57.01649 (7.02)	102.05491 (100), 256.0202 (75.64), 56.01312 (14.96), 212.03032 (9.92), 264.00902 (9.27)
6	D-2-ABA <sup>a</sup>	35.69	133.09636 (100), 61.07526 (35.61), 103.05832 (19.27), 87.05448 (15.64), 287.06172 (15.51)	130.08616 (100), 58.06515 (32.23), 284.05133 (18.09), 84.04437 (11.43), 102.0549 (7.08)
7	L-2-ABA <sup>a</sup>	36.79		
8	DL-proline <sup>c</sup>	36.5	146.09967 (3.26), 72.07187 (0.25), 116.06167 (0.14)	142.08615 (100), 70.0651 (36.39), 114.05493 (18.53), 296.05136 (15.38), 98.0964 (12.39)
9	D-valine <sup>c</sup>	35.97	148.11534 (42.53), 120.08408 (27.61), 103.01746 (21.99)	144.10175 (100), 116.0705 (71.08), 101.0107 (29.15), 98.06 (26.21), 256.0201 (15.78)
10	L-valine <sup>c</sup>	36.9		
11	β-alanine <sup>a</sup>	37.27	273.04611 (10.53), 103.05829	270.03559 (100), 115.06269 (68.9),

			(8.21), 117.06945 (8.21), 101.03375 (7.13), 72.03548 (4.79)	102.05486 (65.84), 98.02357 (57.56), 70.02875 (49.51)
12	3-AIBA <sup>a</sup>	37.87	103.0583 (100), 87.05446 (58.7), 116.05272 (36.36), 162.09454 (21.59), 132.08858 (20.84)	102.05487 (100), 84.04435 (74.81), 112.03922 (55.09), 158.08103 (33.04), 130.04984 (29.1)
13	D-3-ABA <sup>c</sup>	39.15	118.07736 (47.05), 288.06496 (20.26), 72.0355 (15.93)	116.07052 (100), 284.05138 (73.05), 227.02995 (36.09), 70.02878 (34.88), 270.03583 (30.51)
14	L-3-ABA <sup>c</sup>	39.43		
15	isoleucine 1 <sup>d</sup>	39.76	163.13424 (1.86), 103.0175 (1.66)	158.11741 (100), 101.0107 (61.1), 102.05489 (44.75), 283.00722 (28.36), 130.08623 (23.04)
16	isoleucine 2 <sup>d</sup>	40.71		
17	isoleucine 3 <sup>d</sup>	41.04		
18	isoleucine 4 <sup>d</sup>	41.78		
19	D-norvaline <sup>b</sup>	40.49	148.11529 (40.01), 74.08755 (8.48), 258.02713 (5.05), 120.0841 (4.13)	144.10174 (100), 72.08076 (25.05), 256.02012 (13.33), 100.11199 (10.86), 116.07053 (8.62)
20	L-norvaline <sup>b</sup>	41.58		
15	D-leucine <sup>d</sup>	43.11	103.05821 (3.74), 163.13434 (1.84)	158.11737 (100), 102.05486 (92.64), 112.07561 (19.77), 256.02014 (17.91), 101.01071 (12.18)
	L-leucine <sup>d</sup>	44.17		
16	D-norleucine <sup>e</sup>	45.1	163.13427 (9.04)	158.11737 (100), 86.09637 (28.4), 112.07562 (17.19), 102.05489 (14.69), 114.12767 (12.4)
	L-norleucine <sup>e</sup>	46.05		
17	D-aspartic acid <sup>c</sup>	50.52	255.00804 (100), 345.06689 (69.69), 273.04604 (41.25)	254.00453 (100), 342.05658 (60.58), 270.03593 (35.4), 296.01504 (20.64), 228.02527 (17.84)
	L-aspartic acid <sup>c</sup>	50.89		
18	D-serine <sup>b</sup>	56.14	116.0616 (100), 88.03037 (73.53), 103.04934 (29.3), 258.02706 (27.82)	86.0236 (100), 114.05491 (97.06), 256.02023 (27.2), 101.0107 (25.44), 314.02556 (24.72)
	L-serine <sup>b</sup>	56.56		
19	D-2,3-DAPA <sup>b</sup>	72.31	103.05833 (100), 87.04643 (75.88), 159.06752 (45.38), 316.05186 (34.79)	102.05489 (100), 85.03961 (79.61), 157.06066 (44.89), 313.04148 (35.27), 101.01069 (28.47)
20	L-2,3-DAPA <sup>b</sup>	73.2		
21	D-2,4-DABA <sup>d</sup>	76.22	118.07713 (9.88), 131.08076 (7.28)	128.07066 (100), 116.07064 (64.37), 101.01078 (45.24), 129.04208 (28.62), 99.05531 (28.36)
22	L-2,3-DABA <sup>d</sup>	76.79		

512

513 **Table S4** – Intensities of base peaks (obtained by mass extraction of the base peak) for selected amino acids,  
514 according to the three ionization modes (PCI NH<sub>3</sub>, NH<sub>3</sub>/CH<sub>4</sub> and EI at 70 eV), at a concentration of 5 × 10<sup>-5</sup> M.  
515 For sarcosine, proline, leucine and 3-ABA, PCI: molecular ion ((M+H)<sup>+</sup>), except for cysteine (M+H)<sup>+</sup>:  
516 376.0448025).

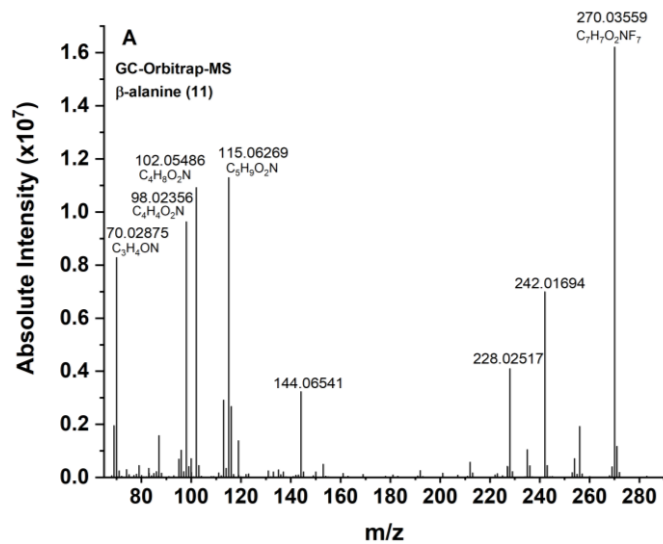
	PCI NH <sub>3</sub>	PCI NH <sub>3</sub> /CH <sub>4</sub>	EI (70eV)
sarcosine	1.79 × 10 <sup>7</sup> (344.07273)	5.53 × 10 <sup>7</sup> (344.07273)	1.35 × 10 <sup>9</sup> (116.07060)
proline	5.05 × 10 <sup>7</sup> (370.08734)	9.77 × 10 <sup>7</sup> (370.08734)	3.90 × 10 <sup>9</sup> (142.08626)
leucine	2.60 × 10 <sup>6</sup> (386.11913)	7.51 × 10 <sup>6</sup> (386.11913)	1.09 × 10 <sup>9</sup> (102.05495)
3-ABA	6.4 × 10 <sup>5</sup> (358.08783)	1.20 × 10 <sup>6</sup> (358.08783)	8.33 × 10 <sup>7</sup> (116.07060)
cysteine	1.13 × 10 <sup>5</sup> (342.05676)	3.02 × 10 <sup>6</sup> (342.05676)	3.47 × 10 <sup>8</sup> (85.98217)

517

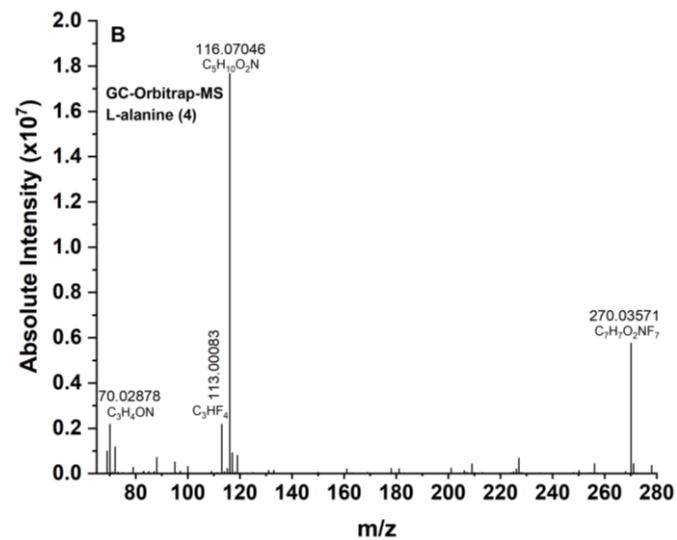
518 **Table S5** - Representative FWHM resolution observed for different fragments of *N*-ethylglycine. The orbitrap  
519 resolution was set at 60 000 at *m/z* 200. The particularity of the orbitrap technology is its higher resolution at low  
520 *m/z*.

<i>N</i> -ethylglycine	
<i>m/z</i>	Mass resolution (FWHM)
130.08621	81367
58.06516	121271
284.05145	53354

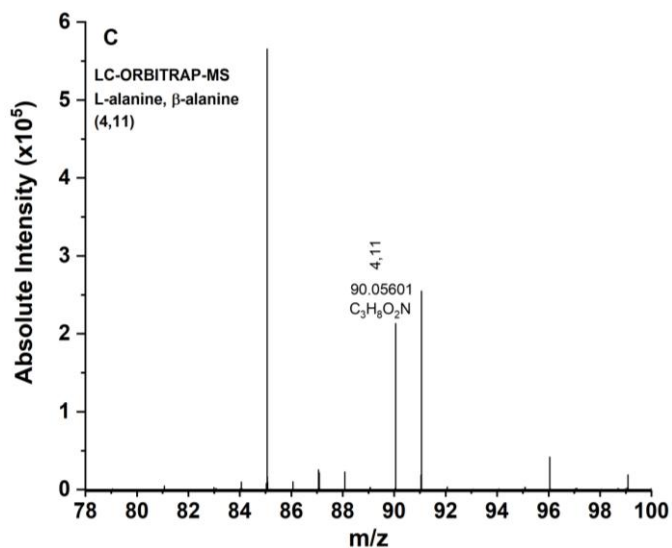
521



522



523



524

525 **Figure S1** – Mass spectra of  $\beta$ -alanine (A) and L-alanine (B) obtained with the electron impact ionization with  
526 GC-FT-orbitrap-MS, compared to the ones obtained with the electro-spray ionization mode with the UPLC-FT-  
527 orbitrap-MS (C). This comparison shows that the EI mode provides additional structural information compared  
528 to the ESI mode. Even with co-elution with the EI mode associated to the HRMS of the orbitrap, compound  
529 identification would be possible.