



**HAL**  
open science

# Identify low mass volatile organic compounds from cometary ice analogs using gas chromatography coupled to an Orbitrap mass spectrometer associated to electron and chemical ionizations

Thomas Javelle, Michel Righezza, Grégoire Danger

## ► To cite this version:

Thomas Javelle, Michel Righezza, Grégoire Danger. Identify low mass volatile organic compounds from cometary ice analogs using gas chromatography coupled to an Orbitrap mass spectrometer associated to electron and chemical ionizations. *Journal of Chromatography A*, 2021, 1652, pp.462343. 10.1016/j.chroma.2021.462343 . hal-03330141

**HAL Id: hal-03330141**

**<https://hal.science/hal-03330141>**

Submitted on 2 Aug 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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Identify Low Mass Volatile Organic Compounds from Cometary Ice Analogs using Gas Chromatography coupled to an Orbitrap mass spectrometer associated to Electron and Chemical Ionizations

Thomas Javelle<sup>1</sup>, Michel Righezza<sup>1</sup> and Grégoire Danger<sup>1,2,3,\*</sup>

<sup>1</sup> Aix-Marseille Université, Laboratoire de Physique des Interactions Ioniques et Moléculaires, UMR 7345, CNRS, Marseille, France

<sup>2</sup> Aix Marseille Université, CNRS, CNES, LAM, Marseille, France

<sup>3</sup> Institut Universitaire de France (IUF)

**ABSTRACT:** Analysis of organic matter extracted from meteorites showed that solar system objects present an important molecular diversity. To improve our understanding of such organic matter, new analytical technologies must be developed. The present study displays the first experiments using a GC-FT-Orbitrap-MS to decipher the molecular diversity observed in experiments simulating the evolution of cometary ices. The proposed analytical strategy focuses on the analysis of 110 volatile organic compounds (VOC) with mainly 1 to 6 carbon atoms generated in such cometary ice analogs. Electron ionization (EI) and chemical ionization (CI) modes with methane (CH<sub>4</sub>) or ammonia (NH<sub>3</sub>) were optimized and compared. Those developments maximized the intensity of molecular, protonated or deprotonated ions and improved high-resolution molecular formula unambiguous identification: EI mode is too energetic to provides there detection, while it is not the case in CI mode. Particularly, NH<sub>3</sub> as a reagent gas improves amine identification in positive mode (PCI) and esters, alcohols, carbonyls, amides, carboxylic acids and nitriles in negative mode (NCI). The combination of both EI and CI mass spectrum analysis improves molecular identification, thanks to the detection of molecular, deprotonated or protonated ion of highest intensity and fragment formula assignments. The EI and NCI NH<sub>3</sub> combination allows formula assignments up to 94 % of our database with limit of detection up to 7 ppm. This procedure has been validated for untargeted GC-FT-Orbitrap-MS analysis of VOC coming from the processing of cometary ice analogs.

**Keywords:** chemical ionization, electron ionization, FT-Orbitrap-MS, GC-MS, high resolution mass spectrometry, volatile organic compound, cometary ice analogs.

## 1. INTRODUCTION

Meteorites are the daughters of interplanetary bodies, comets and asteroids. Their soluble organic compounds were first analysed in 2010 by high resolution mass spectrometry (HRMS) using FT-ICR-MS technology [1]. This analysis within a 150–2,000 Da mass range showed the presence of more than 14,000 organic molecules. Further analytical developments using UPLC-Orbitrap showed that for each detected ion up to 31 different isomers can be present, which suggests the presence of hundreds of thousands of different molecules [2]. To understand the origin of this molecular diversity, laboratory experiments are developed.

In the case of the study of the cometary parent body, the chemistry that led to the organic matter can be simulated in laboratory by starting from ice analogs composed by the primitive and the most abundant compounds (H<sub>2</sub>O, CO, CO<sub>2</sub>,

37 CH<sub>3</sub>OH, NH<sub>3</sub> and H<sub>2</sub>S). Those cometary ice analogs are formed at low temperatures (20 K to 77 K), low pressure (10<sup>-9</sup> to 10<sup>-8</sup> mbar) and then irradiated (ultraviolet photons at Lyman  $\alpha$ , ions or electrons) to simulate their processing when exposed to the surrounding stars or to the Sun irradiation. This alteration leads to the formation of new molecules mainly through the recombination of radicals.

41 To simulate a comet approaching the Sun, the processed cometary ice analogs is then warmed up. It releases many volatile organic compounds (VOC) into the gas phase and a solid organic material remains. Its HRMS analyse (m/z 50-590) revealed the presence of an important molecular diversity with around 10,000 ions with masses up to 4,000 Da [5] presenting similarities with the organic matter observed in meteorites [1,6–8].

45 The composition of the gas phase were still poorly characterized up to 2016 when VOCs was analysed using gas chromatography coupled to mass spectrometry (GC-MS). Thirty-three compounds with m/z up to 102 were identified in the gas phase after the processing and the warming of a cometary ice analog containing only methanol [10]. However, after improvements of the VOC recovery, more than 100 species were still not identified mainly due to the increase of the number of isobars with the rise of molecular masses [9–12]. In this context, the use of HRMS coupled to a gas chromatography will allow the identification of isobars and a better chromatographic reconstruction. It will improve compound identifications and provide information on the chemistry occurring in cometary ice analogs, especially when they contain other heteroatoms like nitrogen (from NH<sub>3</sub>) or sulfur (H<sub>2</sub>S).

53 Invented in the early 2000s [13], the FT-Orbitrap-MS was first coupled with gas chromatography (GC) in 2010 [14] and is only commercially available since 2015 [15]. Nowadays, the FT-Orbitrap-MS is the mass analyser providing the highest resolution when coupled to gas chromatography (120,000 at m/z = 200 Th). The high precision of less than 5 ppm in mass measurement leads to two major enhancements for GC-MS analysis: LOD improvement in full mode thanks to mass deconvolution [15–19], and unambiguous ion formula assignment deduced from ion detection [15–19]. This new technology was already used for pesticide, wastewater pollutants, glyceraldehyde and toxicologic analysis [15–18,20] and for developments aiming to combine mass accuracy to soft ionization to identify unknown [21].

60 In the present work, a GC-FT-Orbitrap-MS is used for the first time to analyse low mass volatile organic compounds formed during experiments simulating the evolution of cometary ices. Conditions involving electron impact ionization (EI) tested at different electron voltage and chemical ionization (CI) with a set of reagent gas are optimized based on standard mixtures. Once the optimal detection of high mass ions was found out, a strategy was developed coupling ionization modes to improve molecular identifications. Limits of detection of various chemical families were also estimated using a semi-quantitative approach. Finally, the identification strategy developed was applied to the analysis of VOC generated by the UV alteration of a cometary ice analogs, demonstrating the interest of using a GC-FT-Orbitrap-MS compared to a classical GC-MS analysis.

## 68 2. MATERIAL AND METHODS

### 69 2.1 Chemicals, database construction and LOD measurement.

70 110 chemicals were purchased from Sigma-Aldrich, Aldrich, Fluka, Fisher, Carlo Ebra, Acros and Riedel-de Haën with the purity of 80 to 99.9 % (see supplementary information TS1 for details). They were chosen according to their astrophysical interest [10,22,23] contains #C = 0 to 6, #O = 0 to 3 and #N = 0 to 1 from 12 different chemical functions (see

73 supplementary information Figure S1). The 500 mg/L standards mixtures methanol (Fisher Chemicals, 99.99 %) solutions  
74 were designed to avoid co-elutions. EI identifications were confirmed with fragment comparison to NIST MAINLIB  
75 database (267376 spectra). CI spectral identification was done using retention time (RT). NIST HR Databases were created  
76 using Deconvolution Plugin 1.3 for TraceFinder 4.1 and NIST MS Search 2.3. Fragment formula are first attribution of  
77 Thermo Scientific Xcalibur 4.1.31.9 within 5 ppm and +1 charge for EI and PCI, and 10 ppm and -1 charge for NCI. LOD  
78 were calculated using dilutions of 40, 8, 4, 0.8 and 0.4 µg/L. LOD are measured on extracted ( $\pm 5$  ppm) most intense peaks  
79 and uncertainty are 95 % Student's t-distribution.

## 80 **2.2 Chemical Synthesis.**

81 Methoxy ethane was synthesized using the Williamson synthesis described by Schmidt et al. [24] using absolute ethanol  
82 (Sigma-Aldrich,  $\geq 99.8$  %), iodomethane (Aldrich) and solid sodium (Aldrich).

## 83 **2.3 Experimental setup.**

84 Data were recorded using a modified GC-FT-Orbitrap-MS (Trace 1310 Gas Chromatograph with an Q-Exactive  
85 Orbitrap<sup>TM</sup> MS analyser from Thermo Fisher Scientific). 1 µL is injected in a GC split/splitless (SSL) injector (250 °C) using  
86 an AI 1310 sampler from Thermo Fisher Scientific. For gas injection, a modified GC SSL injector (280 °C) was used using a  
87 vent-free GC/MS adapter (Frontier Laboratories LTD, 0.15 mm, c.a. 50 cm length, column ultra-ALLOY metal capillarity).  
88 Then the gaseous sample was focused using a MicroJet Cryo-Trap MJT-1035E from Frontier Laboratories LTD  
89 implemented on Rxi® guard column from Restek (20 cm  $\times$  0.25 mm) at 10 cm from the head of the separative column for 45  
90 s using liquid nitrogen (purchased from Linde).

91 GC-FT-Orbitrap-MS has been coupled to a vacuum chamber whose characteristics are detailed elsewhere [4]. Shortly 13.8  
92 mbar of methanol gas (prepared from Fisher Chemicals, 99.99 %) was deposited during 70.2 h on a MgF<sub>2</sub> windows cooled at  
93 77 K using liquid nitrogen (Linde) at low pressure ( $2 \cdot 10^{-7}$  mbar). Concomitantly, the ice was irradiated (mainly Lyman  $\alpha$ )  
94 with a H<sub>2</sub> discharge lamp (121 nm,  $\sim 10^{14}$  photons.cm<sup>-2</sup>.s<sup>-1</sup>). FT-IR transmission spectrometer (BRUKER Vector 22 FTIR  
95 Spectrometer, 4000-910 cm<sup>-1</sup>, 20 scans, 2 cm<sup>-1</sup> resolution) allows to monitor the evolution of the ice while formed and  
96 irradiated. Modifications of the vacuum chamber were based on previously reported studies [9,12]. The photo processed ice  
97 is then warmed up at 300 K and VOC were transferred to the GC-FT-Orbitrap-MS using the previously reported procedure  
98 [9]. The only modification is that both sample injection loops are connected to a unique modified SSL injector.

99 GC analysis was performed using a Rxi®-Sil624MS column (60 m  $\times$  0.25 mm i.d.  $\times$  1.4 µm d.f. from Restek) for all the  
100 analysis. Helium (scientific 6.0 from Linde-France) carrier flow was fixed to 1 mL/min with vacuum compensation,  
101 5.0 mL/min purge flow and split flow of 10.0 mL/min. Temperature into the oven was 35 °C (3 min) up to 130 °C at  
102 2 °C/min to finally reach 280 °C for 2 min at 30 °C/min. Aux Heater 1 and 2 were set at 280 °C and the MS Transfer line at  
103 250 °C.

104 Methane (Alphagaz, Air Liquide) or ammonia (Alphagaz, Air Liquide), in positive (PCI) or negative (NCI) modes were  
105 used. Tuning of PCI CH<sub>4</sub> and NCI CH<sub>4</sub> was done using FT43 and corresponding tune files were used for PCI and NCI NH<sub>3</sub>  
106 analysis. EI source was heated up to 250 °C, PCI CH<sub>4</sub> to 350 °C with 2.0 mL/min gas flow, NCI CH<sub>4</sub> to 200 °C with  
107 1.5 mL/min, PCI NH<sub>3</sub> to 300 °C and 1.5 mL/min and NCI NH<sub>3</sub> to 150 °C with 2.0 mL/min. Except when specified electron  
108 voltage was set at 70 eV (otherwise 10, 30, 50 eV). Data were recorded with resolution  $m/\Delta m = 60\ 000$  at  $m/z = 200$  Th (from

109 142 360 at  $m/z = 45.02146$  Th to 87 795 at  $m/z = 116.08373$  Th, see supplementary information TS2 and TS3). Automated  
110 gain control (AGC) target was 106, maximum precursor ion injection time (ITmax) was 200 ms and maximum IT C-trap  
111 parameters. MS transfer line was set at 250 °C. For gas injections the mass spectrometer recorded within the mass range  
112  $m/z = 30$  to 200 Th, while liquid injection mass range was 40 to 200 Th. All data analyses were monitored using Qual  
113 Browser Thermo Xcalibur 4.1.31.9.

### 114 3. RESULTS

115 The combination of retention time (RT) and low resolution mass spectrometry has been sufficient to identify 33 VOCs  
116 from methanol irradiated ice [10]. Analytical improvement of such analysis shows that many compounds remain unknown  
117 [12]. Extending the database containing such specific compounds is highly expensive and time consuming. In this context,  
118 stoichiometric formula assignments using HRMS improve identification [21,25,26] and optimize standard analysis. The first  
119 step to obtain this information is to optimize high mass ions detection on a set of standard molecules.

#### 120 3.1 Optimization of ionization modes for high mass ion detection.

121 EI is the simplest and most sensitive ionization method used for GC-MS analysis. Unfortunately, this technique produces  
122 unstable excited  $M^{+\bullet}$  radical molecular ions leading to fragmentation. The random detection of those molecular ions prevent  
123 detection of the molecular formula of unknown. A usual way to improve molecular ion detection is to use a “softer”  
124 ionization mode such as CI. In this case proton exchange forms protonated or deprotonated ions which formula are related to  
125 M by the gain or loss of one proton.

##### 126 3.1.1 Electron impact versus chemical ionization global intensity.

127 The low total ion current (TIC) intensity is the principal issue for the use of CI in trace analyses. To investigate the relative  
128 impact of each ionization mode on our targeted molecular families, Figure 1 displays the percentage of TIC for the same  
129 sample for each method (EI at 70 eV, PCI CH<sub>4</sub>, NCI CH<sub>4</sub>, PCI NH<sub>3</sub> and NCI NH<sub>3</sub>) relative to the sum of the total ion current  
130 (TIC) of all methods normalized to 100 for compounds at the same concentration (see supplementary information TS2). It  
131 shows that CI can be more intense than EI for some species. For instance, TIC of NCI NH<sub>3</sub> represents 86 % of the whole  
132 summed TIC of valeraldehyde (64), while TIC of EI at 70 eV, PCI CH<sub>4</sub>, NCI CH<sub>4</sub> or PCI NH<sub>3</sub> corresponds to 7 %, 1 %, 5 %  
133 or 0 % respectively. Consequently, valeraldehyde presents the highest sensitivity with NCI NH<sub>3</sub>. This ionization mode is  
134 optimal for the detection of aldehydes, as well as esters, alcohols, amides or carboxylic acids. On the contrary, EI contributes  
135 to 90 % of the global TIC of 1-methoxypropane (14) versus 2 %, 8 % and 0 % for PCI CH<sub>4</sub>, NCI CH<sub>4</sub> and PCI NH<sub>3</sub>,  
136 respectively. EI at 70 eV is also the optimal ionization mode for the other ethers as well as nitriles and amines. In the case of  
137 these later families, some CI ionization techniques are contributing to more than 5 %. Despite this low TIC intensity, the use  
138 of CI can be considered to obtain interesting structural information: PCI CH<sub>4</sub> for nitriles and amines, PCI NH<sub>3</sub> for amines and  
139 NCI NH<sub>3</sub> for nitriles.

##### 140 3.1.2 Molecular ion in electron impact mode.

141 One possibility to improve the intensity relative to the base peak of the radical molecular ions ( $M^{+\bullet}$ ) in EI is to modulate  
142 the electron energy to limit the amount of energy received by the parent molecule. Indeed, while the total ion formation rate  
143 is optimal at 70 eV, relative abundances of ions with high mass would be higher at 10 eV. When this ion is not detected, the  
144 ideal configuration for molecular formula identification is to detect an ion of high intensity as close as possible to it. In

145 variable electron voltage (VeV) ion resulting from rearrangement like deprotonated cation ( $[M-H]^+$ ) is often detected (see  
146 supplementary information TS3).

147 A standard mixture (see experimental section) was analysed using variable electron voltage from 70 to 10 eV. The  
148 corresponding responses and  $M^{+\bullet}$  or protonated or deprotonated ions detection is displayed in Figure 2. At lower electron  
149 energy, the intensity impacting molecules is weaker allowing to increase the stability of ions with high mass and their  
150 relative abundance without any chemical family discrimination. For instance, from 70 to 10 eV the intensity of the  $M^{+\bullet}$   
151 radical cation of 2-pentanone ( $M^{+\bullet}$ ,  $[C_5H_{10}O]^+$ ,  $m/z = 86.0726$  Th) rises from 9 % to 49 % (Figure 2 a). Thus, for some  
152 analytes, a decrease of the electron energy limits molecule fragmentation and increase the  $M^{+\bullet}$  radical cation intensity.  
153 However, for many other analytes, such as 2-propanol (21,  $M = C_3H_8O$ ), valeric acid (95,  $M = C_5H_{10}O_2$ ) or isopropyl  
154 acetate (49,  $M = C_5H_{10}O_2$ ),  $M^{+\bullet}$  is never detected (TS3) and deprotonated cation relative intensity is also not improved.  
155 Moreover, an important loss of base peak intensity is observed at low electron energy (Figure 2 b). For instance, even with a  
156 significant increase of the relative intensity of molecular or deprotonated cation compared to 70 eV, at 10 eV 2-pentanone  
157 (60) presents a 95 % decrease of the intensity of its base peak. This sensitivity is an issue for analyte identification in trace  
158 analysis. Therefore, the modulation of the electron energy in EI is not relevant to analyse our targeted VOC due to its lack of  
159 selectivity and sensitivity.

### 160 3.1.3 Molecular related ions in chemical ionization modes.

161 To improve the compound identification in CI, one way is to define an ionization mode allowing us to detect an ion of high  
162 intensity as close as possible to the molecular ion. It corresponds to the detection of the protonated or deprotonated cation  
163 ( $[M\pm nH]^+$ ) in PCI or anion ( $[M\pm nH]^-$ ) in NCI of highest intensity. The protonated molecule ( $[M+H]^+$ ,  $n = 1$ , where  $n$   
164 corresponds to the number of hydrogens) in PCI or the deprotonated molecule in NCI ( $[M-H]^-$ ,  $n = 1$ ) may not be detected  
165 because it may be unstable or rearrangements may occur. That is why, we extend this study to deprotonated and protonated  
166 ions from  $n = -5$  up to 5.

167 Figure 3 presents the intensity of molecular, deprotonated or protonated ions of highest intensity relative to the base peak  
168 of a standard molecule grouped according to their chemical families. As shown previously, the detection of molecular ion,  
169 deprotonated or protonated ions of highest intensity in EI mode is difficult due to their fragmentation (Figure 3 a). In general,  
170 NCI modes give the most valuable information on molecular ions, since they allow the detection of either the molecular ion,  
171 or a deprotonated or protonated ion of highest intensity for all families except for ethers (Figure 3 c and e). PCI is much more  
172 selective (Figure 3 b and d). Indeed, PCI  $NH_3$  only provides the detection of one of them for amines, while PCI  $CH_4$  is  
173 selective to esters, amides, carboxylic acids, nitriles and amines. For instance, the detection of one of them of valeraldehyde  
174 (64) is only possible using NCI with  $NH_3$  or  $CH_4$ . PCI  $CH_4$  can only provide its detection for 1-methoxypropane (14) (41 %).  
175 These results confirm that, contrary to variable electron voltage in EI, CI can provide the detection of one of those ions with  
176 a sufficient intensity relative to base peak. However, except for amines with PCI  $NH_3$ , the PCI modes give the lowest TIC  
177 response and it is thus not relevant for our VOC (Figure 2).

178 Once the molecular ion, deprotonated or protonated ions of highest intensity is identified for various analytes, correlations  
179 should be drawn between their chemical structure and their nature. The nature of those ions obtained in CI modes is highly  
180 dependent to their chemical family (see supplementary information TS2). NCI seems to be the most reliable mode. Like in

181 previous studies [21], in NCI, most of them correspond to the deprotonated anions ( $[M-H]^-$ ) for all chemical families except  
182 for ethers and amines. Next to NCI modes, two useful pieces of information can be extracted from PCI modes: amines form  
183 exclusively protonated cation with  $NH_3$  and ethers provide deprotonated cation  $[M-H]^+$  with  $CH_4$ . However, PCI, contrary to  
184 NCI, provides ion of highest mass and highest intensity of different natures for a given family. For instance, in case of  
185 valeraldehyde (64,  $M = C_5H_{10}O$ ) it corresponds to deprotonated cation  $[M-H]^+$  ( $[C_5H_9O]^+$ ,  $m/z = 85.06477$  Th) while 2-  
186 methyl pentanal (80,  $M = C_6H_{12}O$ ) provides protonated cation  $[M+H]^+$ , ( $[C_6H_{13}O]^+$ ,  $m/z = 101.09607$  Th) with PCI  $CH_4$ . In  
187 the same way, valeraldehyde and 2-methyl pentanal provide either the protonated cation  $[M+H]^+$  ( $[C_5H_9O]^+$ ,  
188  $m/z = 85.06473$  Th) and/or the tri-deprotonated cation  $[M-3H]^+$  ( $[C_6H_9O]^+$ ,  $m/z = 97.06470$  Th) with PCI  $NH_3$ , respectively.

189 The NCI  $NH_3$  mode thus produces the highly valuable deprotonated anion in high intensity for esters, alcohols, carbonyls  
190 (aldehydes and ketones), amides, carboxylic acids and nitriles. This technique should give the most valuable molecular  
191 formula information for unknown identification. PCI  $NH_3$  for amines and PCI  $CH_4$  for ethers could be used to obtain the easy  
192 detection of protonated cation in complement to EI for formula identification. For the rest, no CI technique with high TIC  
193 intensity gives molecular ion, deprotonated or protonated ions with sufficient intensity.

### 194 3.2 Semi-quantitative analysis.

195 Limits of detection (LOD) were evaluated for some of our standards (see supplementary information TS1). LOD are  
196 contained in a very wide range between 7 and 4000  $\mu g/L$  (ppm) for nitriles, alcohols, carbonyls and esters. High LOD can be  
197 related to the low sensitivity of gas chromatography with respect to certain chemical families. Amines such as ethylamine (9)  
198 or carboxylic acids like formic acid (37) are usually derivatized before GC-MS analysis [27,28]. For other compounds with  
199 low molecular masses such as ethanol ( $46 \text{ g}\cdot\text{mol}^{-1}$ ) (12) or 1-propanol ( $60 \text{ g/mol}$ ) (32) high LOD are due to detector  
200 limitation as the Orbitrap has inherent low sensitivity around and below  $m/z = 50$  Th. CI and EI LODs are generally in the  
201 same order of magnitude (ratio between 10 and 0.1, Figure S2). Exceptions are alkanes, ethers, nitriles, formic acid (37) and  
202 methyl isobutyrate (56) which are detected more sensitively using EI. Alkanes, ethers and nitriles are often very basic species  
203 that cannot be deprotonated by  $NH_2^-$ , this strong Brønsted base forming deprotonated anion. The lowest LOD are recorded for  
204 4-methyl-2-pentanone (71) within around 7 ppm for EI and  $18 \pm 4 \mu g/L$  (ppm) for NCI  $NH_3$ . This is comparable to analyses  
205 of organics containing 10 C to 13 C in pesticide trace analyses (LOD from 0.5 to 1 ppm - 1  $\mu L$ , splitless injections [15,16]),  
206 toxicology (LOD from 10 to 20 ppm - 2  $\mu L$ , splitless [17]). LOD up to 10 ppb (gas injection, splitless) for wastewater  
207 pollutants detection are only improved thanks to extra preconcentration during sample preparation [18].

## 208 4. DISCUSSION

209 The use of GC-HRMS for the analysis of COV from cometary ice analogs allow molecular formula assignment for  
210 qualitative analysis and enhances LOD thanks to extraction of accurate masses for quantitative analysis. However, HRMS  
211 libraries are not fully available yet. Strategies such as high resolution filtering of low resolution mass spectra hits [29] can be  
212 proceeded, but the comparison of intensity of mass spectra recorded in different conditions is not rigorously possible. Thus,  
213 homemade libraries must be used for compound identifications. The use of soft ionization (CI) for molecular weight  
214 measurement can improve compound identification [25,26]. A common CI method for GC-HRMS identification is PCI  $CH_4$   
215 [26]. However, as shown in this work, in the case of C1 to C6 VOCs analysis PCI  $CH_4$  provides low deprotonated cation  
216 detection with much lower TIC than other CI methods. The combination of EI and NCI  $NH_3$  provides more valuable data for  
217 compound identification with optimal LOD. They allow reliable protonated or deprotonated ion detection, maximal intensity

218 and specific fragmentation features. To improve compound identification, we propose a strategy combining EI and NCI NH<sub>3</sub>  
219 mass spectra for HRMS library search.

#### 220 **4.1 Qualitative analysis.**

221 To exhaustively characterize remaining unknown compounds after database comparison, we propose a workflow  
222 (Figure 4) to combine GC-HRMS analysis from different ionization methods. Step 1 to step 3 are classical steps for  
223 compound identification in GC-MS: acquisition, step 1, mass trace deconvolution and peak detection, step 2, and mass  
224 spectra extraction and library search, step 3. Retention time (RT) alignment, step 4, is a key step to compare different  
225 ionization modes using characteristics found in the present study. We must improve this step because no internal standards  
226 can be used for cometary ice analogs VOCs analysis. Finally, steps 5 and 6 aim to restrict possibilities before standard  
227 analysis confirmation step 7.

##### 228 **4.1.1 Simple ionization mode GC-HRMS analysis.**

229 After recording the chromatographic data using different ionization sources, step 1, a list of detected compounds with their  
230 corresponding identification is separately edited: chromatographic peaks are listed thanks to peak detection on narrow  
231 ( $\pm 5$  ppm) extraction of mass traces, step 2. The corresponding mass spectra and RT are compared with the existing database.  
232 Mass spectrum comparison is obtained thanks to the mass spectrum similarity index defined as

$$233 \quad SI = \left( 1 - \frac{\sum_{m/z} |I_u(m/z) - I_l(m/z)|}{\sum_{m/z} |I_u(m/z) + I_l(m/z)|} \right) \times 100 \quad \text{Eq. 1 [30]},$$

234 with  $I_u(m/z)$  the intensity of the unknown at exact mass  $m/z$  and  $I_l(m/z)$ , the intensity of the compound referenced in the  
235 library at same exact  $m/z$ , and the RT difference.

236 As shown previously using reference spectra, NCI NH<sub>3</sub> is used for esters, alcohols, carbonyls, amides, carboxylic acids  
237 and nitriles analysis. During unknown analysis of such compounds, deprotonated anion can be identified as an ion of highest  
238 mass with intensity relative to base peak higher than 2.5 %. This is true for 94 % of analysed standards (see supplementary  
239 material Table S1). If the second highest mass ion is considered, this ion is included for 94 % of them. The remaining 6 %  
240 contain trimethylacetaldehyde (35) that forms adduct induced by hyperconjugation of its tert butyl group, methyl formate (6)  
241 and ethyl formate (19) that are the smaller esters and it still present but at lower threshold (1.19 and 1.60 resp., see  
242 supplementary material Table S5). From its detection, direct measurement of the number of #O, #N, #C and #H and DBE can  
243 help to characterize unknowns.

244 Furthermore, NCI NH<sub>3</sub> can give information on primary neutral fragments helping to characterize chemical structures by  
245 allowing discriminations between different chemical families. The presence of a fragment with a mass loss of 16.03132 Th (1  
246 C + 4 H) relative to the deprotonated anion is characteristic to compounds with a vicinal methyl (alcohols, carbonyls or  
247 nitriles). This mass loss is consistent with alcohol groups in position 2. For example, 1-pentanol (83) and 3-pentanol (67) do  
248 not produce such loss, while 2-pentanol (68) does. 2-methyl-2-butanol (48), 2-propanol (21) and 3-methyl-2-butanol (61)  
249 follow the same rule. This mass loss is also recorded for 2-ketones. Indeed, 2-pentanone (60) provides it but not 3-pentanone  
250 (65). This rule is consistent with acetone (18) and 4-methyl-2-pentanone (71) fragmentation patterns. Finally, methyl groups  
251 in  $\beta$  or  $\gamma$ -position of aldehydes or nitriles can also be lost. This is the case for isovaleraldehyde (53), isovaleronitrile (78) and  
252 2-methyl butyronitrile (72). Isobutyraldehyde (27) is the only unexplained exception. These observations are useful to

253 understand most of the NCI NH<sub>3</sub> fragmentation patterns observed for our standards and to restrict possible structures. The  
254 protonated or deprotonated ion detection (using PCI NH<sub>3</sub> for amines and NCI NH<sub>3</sub> for the others) and identification of mass  
255 loss with NCI NH<sub>3</sub> can therefore give information on possible structures. Once the formula of the molecular ion and its  
256 branching are known, the last unknown information is the chemical family. This can be obtained by matching information  
257 from multiple ionization modes.

#### 258 **4.1.2 Multiple ionization mode GC-HRMS analysis.**

259 Combining information from multiple ionization modes requires the possibility to match peak lists containing both  
260 identified and remaining unknown from two GC-HRMS analysis. The step allowing to match lists is known as RT  
261 alignment, step 4. Indeed, a direct match is not possible because of the usual RT shift between two recorded data. RT can be  
262 corrected either using the retention index relative to an internal standard [31] or by cross comparison of RT and mass spectra  
263 [32]. Unfortunately, none of those two usual techniques can be used in the context of GC-HRMS analysis of ice analogs with  
264 multiple ionization modes. Internal standards cannot be used during ice analog synthesis or added before VOC analyses.  
265 Furthermore, mass spectra and retention times cannot be fitted to align RT because of the different ionization modes.  
266 Consequently, we chose to align RT by matching identified compounds and using a narrow RT uncertainty window  $\pm 1$  %  
267 for unknown peaks. The matched list can then be used to characterize unknown compounds combining the different  
268 ionization sources properties.

269 An exhaustive analysis of normalized TIC intensity, step 5, combining different ionization methods allows to sort  
270 unknowns into four distinct groups, alkanes/ethers, amines, nitriles and others (esters, alcohols, aldehydes, amides and  
271 carboxylic acids), using two steps (Table 1). First, the comparison of EI to NCI NH<sub>3</sub> classifies molecules into three groups  
272 according to the normalized TIC contribution of NCI NH<sub>3</sub>: high, medium and low (Figure 1). One of them contains esters,  
273 alcohols, aldehydes, amides and carboxylic acids having high NCI NH<sub>3</sub> normalized TIC contribution: ~93, 87, ~92, 98 or  
274 92 %, respectively. The other only nitriles (32 % of NCI NH<sub>3</sub> contribution). The last one corresponds to alkanes, ethers and  
275 amines with low NCI NH<sub>3</sub> contribution (0, ~7 and ~4 % resp.). Second, extra comparison of PCI NH<sub>3</sub> with EI can isolate  
276 amines: the contribution of PCI NH<sub>3</sub> relative to EI for such molecules is around 33 % for tested compounds (Figure 1).

277 The combination of the chemical group with HRMS spectral data, step 6, allows more precise unknown identification,  
278 thanks to formula identification (6a) and mass spectra analysis (6b). Step (6a) can split the groups alkanes/ethers and others  
279 (Table 1). The presence of an oxygen atom into the chemical formula of an unknown contained into the group alkane/ethers  
280 proves it is an ether. Using the same arguments one can split the chemical group others into amides ( $\#O = 1$  and  $\#N = 1$ ),  
281 esters/carboxylic acids ( $\#O = 2$  and  $\#N = 0$ ) and carbonyls ( $\#O = 1$  and  $\#N = 0$ ). At this step, seven groups exist: alkanes,  
282 ethers, amines, nitriles, amides, esters/carboxylic acids and carbonyls (Table 1).

283 During mass spectra analysis (6b), primary mass loss must be considered. Indeed, the mass loss of 2.01557 Th (2 H) with  
284 respect to the deprotonated anion is interesting to distinguish alcohols from carbonyls mass spectra. The formation of this ion  
285 is highly efficient for alcohols and corresponds to their base peak ( $DBE = 1.5$  for saturated alcohols) while the deprotonated  
286 anion ( $DBE = 0.5$  in case of saturated alcohols) corresponds to a much lower intensity peak (from 3 to 10 % of intensity).  
287 This way, 2-methyl-1-propanol (44) can be distinguished from isobutyraldehyde (27) because it produces such ions very  
288 efficiently (100 %) while only traces (0.64 %) for isobutyraldehyde (see supplementary information Table S4).

289 Finally, mass spectra of esters and carboxylic acids can be distinguished thanks to mass spectra analysis. The presence of  
290 any fragment is characteristic to esters because of the breakable C(=O)-O-R bound (see supplementary information Table  
291 S5) while carboxylic acids do not provide any fragment (only adducts sometimes, see supplementary information Table S6)  
292 using NCI NH<sub>3</sub>. These elements show that the combination of multiple ionization modes can provide a full characterization  
293 of an unknown compound. Moreover, as previously described, branching information can also be described using NCI NH<sub>3</sub>.  
294 Finally for ethers and alkanes that are the only molecular families undetectable using CI, exhaustive analysis of EI mass  
295 spectra can help to further interpretation and enable standard proposition to definitive identification.

#### 296 **4.2 Application to VOCs identification from cometary ice analogue.**

297 In part 4.1, we proposed a workflow combining different ionization modes to develop an analytical strategy for an  
298 exhaustive identification of VOC. In this section, we test this strategy to identify VOC formed from a methanol ice  
299 previously irradiated at Lyman  $\alpha$  and at low temperature (77 K). Once irradiated, this ice is warmed up to 300 K and the  
300 VOCs desorbing during this warming are recovered to be analyzed with the GC-HRMS (details in methods section).

##### 301 **4.2.1 Standard identification.**

302 The classical identification process was monitored using the database (see supplementary information Table S12) with the  
303 *SI* (Eq. 1) indicator for high resolution mass spectra ( $\Delta m/m < 5$  ppm) and retention time comparison. This allowed the  
304 identification of dioxygen (2), carbon dioxide (3), ether (4), acetaldehyde (5), methanol (8), methyl formate (6), ethanol (12),  
305 diethyl ether (11) and methyl acetate (22) (Figure 5). Identifications are commonly confirmed using both EI and NCI NH<sub>3</sub>  
306 ionization methods. The only exception is dioxygen (2) that has been only detected by EI because of its better sensitivity, and  
307 methyl acetate (22) because NCI NH<sub>3</sub> has better sensitivity for esters than EI.

308 Monitoring both EI and NCI NH<sub>3</sub> analysis enables deconvolution of coeluting species while extracted ion chromatograms  
309 cannot distinguish them. For instance, Peak@9.31 with a NCI NH<sub>3</sub> normalized TIC relative to EI of 60 % can be identified  
310 as diethyl ether using EI (11,  $\Delta RT = -0.04$  min, *SI* = 80 %) coeluting with ethanol using NCI NH<sub>3</sub> (12,  $\Delta RT = -0.16$  min,  
311 *SI* = 48 %). Again, this can be explained by sensitivity issues: NCI NH<sub>3</sub> has lower sensitivity than EI for ethers, while it is  
312 the contrary for alcohols (Figure 1).

##### 313 **4.2.2 Cross-validation.**

314 To cross-validate the strategy, dimethyl ether (4) and acetaldehyde (5) were removed from the mass spectra library  
315 (Table 2). Consequently, after steps 1 to 4 (identification and mass extraction, Figure 4), two extra unidentified peaks were  
316 present in both chromatograms: Peak@5.81 and Peak@6.90. The description of the multiple ionization modes showed that  
317 for the peak at 5.81 min the normalized TIC of NCI NH<sub>3</sub> relative to EI was 2 %. This unknown is part of “low normalized  
318 TIC contribution of NCI NH<sub>3</sub>” i.e. alkanes, ethers and amines. However, as the chemical formula of its NCI NH<sub>3</sub> major  
319 fragment (CH<sub>3</sub>O, *DBE* = 0.5, see supplementary information Table S10) contains one heteroatom (*#O* = 1 and *#N* = 0),  
320 Peak@5.81 must be an ether. As discussed previously, ethers do not provide deprotonated anions using NCI NH<sub>3</sub>, but the EI  
321 mode can help to identify it. Indeed, the only EI fragments with intensity higher than 2.5 % are [C<sub>2</sub>H<sub>5</sub>O]<sup>+</sup>  
322 (*m/z* = 45.03350 Th, *DBE* = 0.5) and [C<sub>2</sub>H<sub>6</sub>O]<sup>+</sup> (*m/z* = 46.04130 Th, *DBE* = 0, see supplementary information Table S11).  
323 [C<sub>2</sub>H<sub>6</sub>O]<sup>+</sup> is the ion with the highest mass having a *DBE* as an entire. In EI mass spectrometry, such clues show that it could  
324 be the M<sup>+</sup> radical cation [33]. Consequently, Peak@5.81 can be an ether with *M* = C<sub>2</sub>H<sub>6</sub>O, i.e. dimethyl ether.

325 The Peak@6.90 contains two fragments in its NCI NH<sub>3</sub> spectrum with intensity higher than 2.5 %: [C<sub>2</sub>HO]<sup>-</sup>  
326 ( $m/z = 41.00352$  Th,  $DBE = 2.5$ ) and [C<sub>2</sub>H<sub>3</sub>O]<sup>-</sup> ( $m/z = 43.01918$  Th,  $DBE = 1.5$ ). As discussed previously, this ion of highest  
327 mass and highest intensity corresponds to the deprotonated anion on NCI NH<sub>3</sub> mass spectra. Consequently, C<sub>2</sub>H<sub>3</sub>O = M-H  
328 i.e. its molecular formula is C<sub>2</sub>H<sub>4</sub>O. Peak@6.90 is consequently an unsaturated molecule ( $DBE = 1$ ). To improve information  
329 upon this unknown, EI and NCI NH<sub>3</sub> normalized TIC response can be compared. This molecule is part of the “high  
330 normalized TIC contribution of NCI NH<sub>3</sub>” (90 %). Its chemical family can correspond to esters, alcohols, aldehydes, amides  
331 or carboxylic acids. While for Peak@6.90 #O = 1 and #N = 0, one can directly exclude esters (#O = 2), amides (#N = 1) and  
332 carboxylic acids (#O = 2) from the last list. The two remaining families are carbonyls or alcohols. To argue for one or the  
333 other, one must look for the relative intensity of the mass loss of 2.01566 Th (2 H) with respect to ion of highest mass  
334 ([C<sub>2</sub>HO]<sup>-</sup>, 4 %) versus the deprotonated anion (100 %). The fact that the intensity of the second one is higher than this mass  
335 loss indicates that this chromatographic peak is not alcohol but carbonyls. Consequently, Peak@6.90 is acetaldehyde. These  
336 identifications are consistent with the previous.

337 Finally, using the same strategy, we propose an assignment to the five remaining unknown that has not been identified  
338 previously (Table 2). Two must be removed, as they are peaks containing ions from the column and are present into blanks.  
339 Then, from the five remaining unknowns, Peak@6.74 contains the same ions as acetaldehyde. It is a double injection peak,  
340 certainly due to slow heating of the cryotrapping. The last unidentified peaks are only detected using the EI mode. They may  
341 be out of detection limits obtained with NCI NH<sub>3</sub>. Peak@5.30, Peak@7.22 and Peak@9.01 can be part of groups alkanes,  
342 ethers, amines or nitriles (Table 1). Peak@5.30 EI mass spectrum is consistent with C<sub>3</sub> alkane formation (propane or  
343 propene): [C<sub>3</sub>H<sub>n</sub>]<sup>+</sup> the major peaks with n from 2 to 7. Comparison with standard mass spectrum shows that Peak@5.30  
344 cannot be propane. Peak@7.22's ion of highest mass is [C<sub>3</sub>H<sub>7</sub>O]<sup>+</sup> ( $m/z = 59.04922$  Th,  $DBE = 0.5$ ). Methoxy ethane is the  
345 only unsaturated ether with this molecular mass. This attribution has been confirmed with the presence of masses  
346 45.03347 Th and 59.04920 Th at the right RT after comparison with the analysis of the products of a Williamson synthesis.  
347 Finally, the only ion present in the EI-MS of Peak@9.01 is [C<sub>3</sub>H<sub>7</sub>O]<sup>+</sup> ( $m/z = 59.04922$  Th,  $DBE = 0.5$ ). This implies that it is  
348 an ether with  $M = C_3H_7O-R$  with unknown alkyl chain R with fragment 59 which must be more intense than fragments 31,  
349 43 or 49. Comparing RT with the one recorded in our database and the intensity of fragment 59 intensity relative to the three  
350 others fragments in mass spectra from the NIST database 2-methoxy propane seems to be the most probable identification of  
351 Peak@9.01.

## 352 5. CONCLUSION

353 In the present study, a GC-FT-Orbitrap-MS was used for the first time to analyse VOCs containing 1 to 6 carbon atoms  
354 produced during the photo-processing of cometary ice analogs. The recovery of these VOCs was allowed by the specific  
355 coupling of the GC-FT-Orbitrap-MS to the vacuum chamber where ices are processed. Different ionization modes (EI, PCI or  
356 NCI with NH<sub>3</sub> or CH<sub>4</sub>) were first tested. Their combination with the FT-Orbitrap-MS high-resolution (full width at half  
357 maximum over 100.000 at  $m/z = 200$ ) and precision (less than 5ppm) provides unambiguous formula identifications and  
358 untargeted identification of molecular ions. Variable electron voltage ionization methods cannot provide molecular ion  
359 improvement without drastic loss of TIC intensity. With chemical ionization modes (methane or ammonia in positive or  
360 negative ionization modes), we demonstrate that NCI NH<sub>3</sub>, generally poorly used, is the best mode to use for compounds  
361 having less than six carbon atoms. The best compromise for sensitivity and easy molecular formula identification consists in

362 combining EI with NCI NH<sub>3</sub>, for all compounds, except amines, alkanes and ethers, with an accuracy of formula detection of  
363 94 %. Semi-quantitative analysis shows that NCI NH<sub>3</sub> and EI have similar LODs up to 7 ppm (database of 110 compounds  
364 containing 1 to 6 carbon atoms Table S1). In the case of amine identification, the ideal combination is EI with PCI NH<sub>3</sub>,  
365 while only EI can ionize alkanes and ethers.

366 From those results, a workflow (Figure 4) has been built up using formula identification thanks to high resolution mass  
367 spectra and chromatographic normalized TIC comparison of CI with respect to EI. This workflow has been cross-validated  
368 on the identification of unknown VOCs generated by the photo-processing of an ice analog of methanol. Specially, dimethyl  
369 ether (4), acetaldehyde (5) and methoxy ethane were identified and validated while propene and 2-methoxy propane were  
370 identified without standard analysis validation. Further improvements of this method will be to identify compounds from  
371 nitrogen containing cometary ices analogs.

372

## 373 **ASSOCIATED CONTENT**

### 374 **Supporting material**

375 Figure S1. Composition of the library.

376 Figure S2. EI over NCI LOD ratio.

377 Figure S3. Summary of the relative intensity of the ions of highest mass and intensity related to the chemical formula for  
378 NCI NH<sub>3</sub> ionization of the total database analysis.

379 Figure S4. Infrared spectrum of CH<sub>3</sub>OH after 2.7h irradiation while deposition.

380 Table S1. Calculated and measured data for the 110 analytes included in the database.

381 Table S2. Summary of EI/CI ion of highest mass and intensity related to the chemical formula.

382 Table S3. Summary of VeV ion of highest mass and intensity related to the chemical formula.

383 Table S4. Comparison of intensity of ion of highest mass and intensity related to the chemical formula detected in NCI  
384 NH<sub>3</sub> for carbonyls and alcohols.

385 Table S5. Comparison of NCI NH<sub>3</sub> detected ions intensity for esters.

386 Table S6. Comparison of NCI NH<sub>3</sub> detected ions intensity for carboxylic acids.

387 Table S7. Comparison of NCI NH<sub>3</sub> detected ions intensity for ethers.

388 Table S8. Comparison of NCI NH<sub>3</sub> detected ions intensity for amines.

389 Table S9. Comparison of NCI NH<sub>3</sub> detected ions intensity for nitriles.

390 Table S10. NCI NH<sub>3</sub> extracted unknown mass spectra.

391 Table S11. EI extracted unknown mass spectra.

## 392 **AUTHOR INFORMATION**

### 393 **Corresponding Author**

394 \*E-mail: [gregoire.danger@univ-amu.fr](mailto:gregoire.danger@univ-amu.fr)

### 395 **Author Contributions**

396 TJ performed the experiments, treated the data and developed the methodology. TJ, GD and MR wrote the manuscript and

397 GD managed the program.

398

#### 399 **Notes**

400 The authors declare no competing financial interest.

401

#### 402 **ACKNOWLEDGMENT**

403 G.D. and T.J. are highly grateful to Y. Guillaneuf and C. Lefay from Institut de Chimie Radicalaire, UMR CNRS 7273 for  
404 their help for the synthesis of methoxy ethane. G.D. thank the Région Sud/Provence-Alpes-Côte d'Azur, the Agence  
405 Nationale de la Recherche (RAHIA\_SSOM, ANR-16-CE29-0015 and VAHIA, ANR- 12-JS08-0001), the Centre National  
406 d'Etudes Spatiales from its exobiology program, and the Centre National de la Recherche Française (CNRS, "Physique et  
407 Chimie du Milieu Interstellaire" (PCMI) and "Programme National de Planétologie" (PNP) programs) for their financial  
408 support. T.J. thank ENS Lyon for foundings.

#### 409 **REFERENCES**

- 410 [1] P. Schmitt-Kopplin, Z. Gabelica, R.D. Gougeon, A. Fekete, B. Kanawati, M. Harir, I. Gebefuegi, G. Eckel, N.  
411 Hertkorn, High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall,  
412 Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 2763–2768, doi: 10.1073/pnas.0912157107.
- 413 [2] A. Ruf, P. Poinot, C. Geffroy, L. Le Sergeant d'Hendecourt, G. Danger, Data-Driven UPLC-Orbitrap MS Analysis  
414 in Astrochemistry, Life. 9 (2019), doi: 10.3390/life9020035.
- 415 [3] S.A. Sandford, M. Nuevo, P.P. Bera, T.J. Lee, Prebiotic Astrochemistry and the Formation of Molecules of  
416 Astrobiological Interest in Interstellar Clouds and Protostellar Disks, Chem. Rev. 120 (2020) 4616–4659, doi:  
417 10.1021/acs.chemrev.9b00560.
- 418 [4] L. d'Hendecourt, P. Modica, C. Meinert, L. Nahon, U. Meierhenrich, Interstellar ices: a possible scenario for  
419 symmetry breaking of extraterrestrial chiral organic molecules of prebiotic interest, Journal of Interdisciplinary  
420 Methodologies and Issues in Science. (2019). <https://hal.archives-ouvertes.fr/hal-01842407> (accessed May 15, 2019).
- 421 [5] G. Danger, F.-R. Orthous-Daunay, P. de Marcellus, P. Modica, V. Vuitton, F. Duvernay, L. Flandinet, L. Le  
422 Sergeant d'Hendecourt, R. Thissen, T. Chiavassa, Characterization of laboratory analogs of interstellar/cometary organic  
423 residues using very high resolution mass spectrometry, Geochim. Cosmochim. Acta. 118 (2013) 184–201, doi:  
424 10.1016/j.gca.2013.05.015.
- 425 [6] A. Fresneau, N. Abou Mrad, L. LS d'Hendecourt, F. Duvernay, L. Flandinet, F.-R. Orthous-Daunay, V. Vuitton, R.  
426 Thissen, T. Chiavassa, G. Danger, Cometary Materials Originating from Interstellar Ices: Clues from Laboratory  
427 Experiments, The Astrophysical Journal. 837 (2017) 168, doi: 10.3847/1538-4357/aa618a.
- 428 [7] R. Briggs, G. Ertem, J.P. Ferris, J.M. Greenberg, P.J. McCain, C.X. Mendoza-Gomez, W. Schutte, Comet Halley as  
429 an aggregate of interstellar dust and further evidence for the photochemical formation of organics in the interstellar medium,  
430 Orig. Life Evol. Biosph. 22 (1992) 287–307, doi: 10.1007/BF01810858.
- 431 [8] Á. Somogyi, R. Thissen, F.-R. Orthous-Daunay, V. Vuitton, The Role of Ultrahigh Resolution Fourier Transform  
432 Mass Spectrometry (FT-MS) in Astrobiology-Related Research: Analysis of Meteorites and Tholins, Int. J. Mol. Sci. 17  
433 (2016) 439, doi: 10.3390/ijms17040439.

- 434 [9] N. Abou Mrad, F. Duvernay, P. Theulé, T. Chiavassa, G. Danger, Development and optimization of an analytical  
435 system for volatile organic compound analysis coming from the heating of interstellar/cometary ice analogues, *Anal. Chem.*  
436 86 (2014) 8391–8399, doi: 10.1021/ac501974c.
- 437 [10] N. Abou Mrad, F. Duvernay, T. Chiavassa, G. Danger, Methanol ice VUV photoprocessing: GC-MS analysis of  
438 volatile organic compounds, *Mon. Not. R. Astron. Soc.* 458 (2016) 1234–1241, doi: 10.1093/mnras/stw346.
- 439 [11] N. Abou Mrad, F. Duvernay, R. Isnard, T. Chiavassa, G. Danger, The Gaseous Phase as a Probe of the  
440 Astrophysical Solid Phase Chemistry, *ApJ.* 846 (2017) 124, doi: 10.3847/1538-4357/aa7cf0.
- 441 [12] N. Abou Mrad, S. Werner, J. Mouzay, G. Danger, Sensitivity and resolution optimization in gas chromatography  
442 coupled to mass spectrometry analyses of volatile organic compounds present in vacuum environment, *J. Chromatogr. A.*  
443 1609 (2020) 460489, doi: 10.1016/j.chroma.2019.460489.
- 444 [13] A. Makarov, Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis,  
445 *Anal. Chem.* 72 (2000) 1156–1162, doi: 10.1021/ac991131p.
- 446 [14] A.C. Peterson, G.C. McAlister, S.T. Quarmby, J. Griep-Raming, J.J. Coon, Development and characterization of a  
447 GC-enabled QLT-Orbitrap for high-resolution and high-mass accuracy GC/MS, *Anal. Chem.* 82 (2010) 8618–8628, doi:  
448 10.1021/ac101757m.
- 449 [15] H.G.J. Mol, M. Tienstra, P. Zomer, Evaluation of gas chromatography - electron ionization - full scan high  
450 resolution Orbitrap mass spectrometry for pesticide residue analysis, *Anal. Chim. Acta.* 935 (2016) 161–172, doi:  
451 10.1016/j.aca.2016.06.017.
- 452 [16] M. Vargas-Pérez, I. Domínguez, F.J.E. González, A.G. Frenich, Application of full scan gas chromatography high  
453 resolution mass spectrometry data to quantify targeted-pesticide residues and to screen for additional substances of concern  
454 in fresh-food commodities, *J. Chromatogr. A.* 1622 (2020) 461118, doi: 10.1016/j.chroma.2020.461118.
- 455 [17] A. Rivera-Pérez, R. López-Ruiz, R. Romero-González, A. Garrido Frenich, A new strategy based on gas  
456 chromatography-high resolution mass spectrometry (GC-HRMS-Q-Orbitrap) for the determination of alkenylbenzenes in  
457 pepper and its varieties, *Food Chem.* 321 (2020) 126727, doi: 10.1016/j.foodchem.2020.126727.
- 458 [18] I. Domínguez, F.J. Arrebola, J.L. Martínez Vidal, A. Garrido Frenich, Assessment of wastewater pollution by gas  
459 chromatography and high resolution Orbitrap mass spectrometry, *J. Chromatogr. A.* 1619 (2020) 460964, doi:  
460 10.1016/j.chroma.2020.460964.
- 461 [19] A. Gujar, J. Voss, J. Cole, X. Zheng, D. Cavagnino, Deconvolution in TraceFinder software for unitresolution GC-  
462 MS data, ThermoFisher Scientific. (2018). [https://assets.thermofisher.com/TFS-Assets/CMD/Technical-Notes/tn-10624-gc-  
463 ms-deconvolution-plugin-tracefinder-tn10624-en.pdf](https://assets.thermofisher.com/TFS-Assets/CMD/Technical-Notes/tn-10624-gc-ms-deconvolution-plugin-tracefinder-tn10624-en.pdf) (accessed May 28, 2020).
- 464 [20] S. Fathalinejad, E. Taarning, P. Christensen, J.H. Christensen, The development and validation of a GC-MS method  
465 for the quantification of glycolaldehyde formed from carbohydrate fragmentation processes, *Anal. Methods.* 12 (2020) 1975–  
466 1987, doi: 10.1039/C9AY02639H.
- 467 [21] B.B. Misra, M. Olivier, High Resolution GC-Orbitrap-MS Metabolomics Using Both Electron Ionization and  
468 Chemical Ionization for Analysis of Human Plasma, *J. Proteome Res.* (2020), doi: 10.1021/acs.jproteome.9b00774.
- 469 [22] C. Szopa, R. Sternberg, F. Raulin, H. Rosenbauer, What can we expect from the in situ chemical investigation of a  
470 cometary nucleus by gas chromatography: First results from laboratory studies, *Planetary and Space Science.* 51 (2003) 863–

471 877, doi: 10.1016/s0032-0633(03)00118-1.

472 [23] M.F. A'Hearn, L.M. Feaga, H. Uwe Keller, H. Kawakita, D.L. Hampton, J. Kissel, K.P. Klaasen, L.A. McFadden,  
473 K.J. Meech, P.H. Schultz, J.M. Sunshine, P.C. Thomas, J. Veverka, D.K. Yeomans, S. Besse, D. Bodewits, T.L. Farnham, O.  
474 Groussin, M.S. Kelley, C.M. Lisse, F. Merlin, S. Protopapa, D.D. Wellnitz, COMETARY VOLATILES AND THE ORIGIN  
475 OF COMETS, *The Astrophysical Journal*. 758 (2012) 29, doi: 10.1088/0004-637x/758/1/29.

476 [24] F. Schmidt, P. Swiderek, J.H. Bredehöft, Electron-Induced Formation of Ethyl Methyl Ether in Condensed Mixtures  
477 of Methanol and Ethylene, *J. Phys. Chem. A*. 123 (2019) 37–47 10, doi: 10.1021/acs.jpca.8b10209.

478 [25] M. Ubukata, A. Kubo, K. Nagatomo, T. Hizume, H. Ishioka, A.J. Dane, R.B. Cody, Y. Ueda, Integrated qualitative  
479 analysis of polymer sample by pyrolysis-gas chromatography combined with high-resolution mass spectrometry: Using  
480 accurate mass measurement results from both electron ionization and soft ionization, *Rapid Commun. Mass Spectrom.* 34  
481 (2020) e8820, doi: 10.1002/rcm.8820.

482 [26] S. Abate, Y.G. Ahn, T. Kind, T.R.I. Cataldi, O. Fiehn, Determination of elemental compositions by gas  
483 chromatography/time-of-flight mass spectrometry using chemical and electron ionization, *Rapid Commun. Mass Spectrom.*  
484 24 (2010) 1172–1180, doi: 10.1002/rcm.4482.

485 [27] F. Sacher, S. Lenz, H.-J. Brauch, Analysis of primary and secondary aliphatic amines in waste water and surface  
486 water by gas chromatography-mass spectrometry after derivatization with 2,4-dinitrofluorobenzene or benzenesulfonyl  
487 chloride, *J. Chromatogr. A*. 764 (1997) 85–93, doi: 10.1016/S0021-9673(96)00868-0.

488 [28] Y. Sha, J. Meng, Y. Zhang, C. Deng, D. Wu, Determination of volatile organic acids in tobacco by single-drop  
489 microextraction with in-syringe derivatization followed by GC-MS, *J. Sep. Sci.* 33 (2010) 212–217, doi:  
490 10.1002/jssc.200900523.

491 [29] N.W. Kwiecien, D.J. Bailey, M.J.P. Rush, J.S. Cole, A. Ulbrich, A.S. Hebert, M.S. Westphall, J.J. Coon, High-  
492 resolution filtering for improved small molecule identification via GC/MS, *Anal. Chem.* 87 (2015) 8328–8335, doi:  
493 10.1021/acs.analchem.5b01503.

494 [30] C.R. Warren, Use of chemical ionization for GC-MS metabolite profiling, *Metabolomics*. 9 (2013) 110–120, doi:  
495 10.1007/s11306-011-0346-8.

496 [31] J.S. Lübeck, G.L. Alexandrino, J.H. Christensen, GC × GC-HRMS nontarget fingerprinting of organic  
497 micropollutants in urban freshwater sediments, *Environ Sci Eur*. 32 (2020) 1072, doi: 10.1186/s12302-020-00353-2.

498 [32] T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: modular framework for processing, visualizing,  
499 and analyzing mass spectrometry-based molecular profile data, *BMC Bioinformatics*. 11 (2010) 395, doi: 10.1186/1471-  
500 2105-11-395.

501 [33] R. Venkataraghavan, F.W. McLafferty, G.E. van Lear, Computer-aided interpretation of mass spectra, *Org. Mass*  
502 *Spectrom.* 2 (1969) 1–15, doi: 10.1002/oms.1210020102.

503  
504

# Identify Low Mass Volatile Organic Compounds from Cometary Ice Analogs using Gas Chromatography coupled to an Orbitrap mass spectrometer associated to Electron and Chemical Ionizations

Thomas Javelle<sup>1</sup>, Michel Righezza<sup>1</sup> and Grégoire Danger<sup>1,2,3,\*</sup>

<sup>1</sup> Aix-Marseille Université, Laboratoire de Physique des Interactions Ioniques et Moléculaires, UMR 7345, CNRS, Marseille, France

<sup>2</sup> Aix Marseille Université, CNRS, CNES, LAM, Marseille, France

<sup>3</sup> Institut Universitaire de France (IUF)

## Figures and Tables

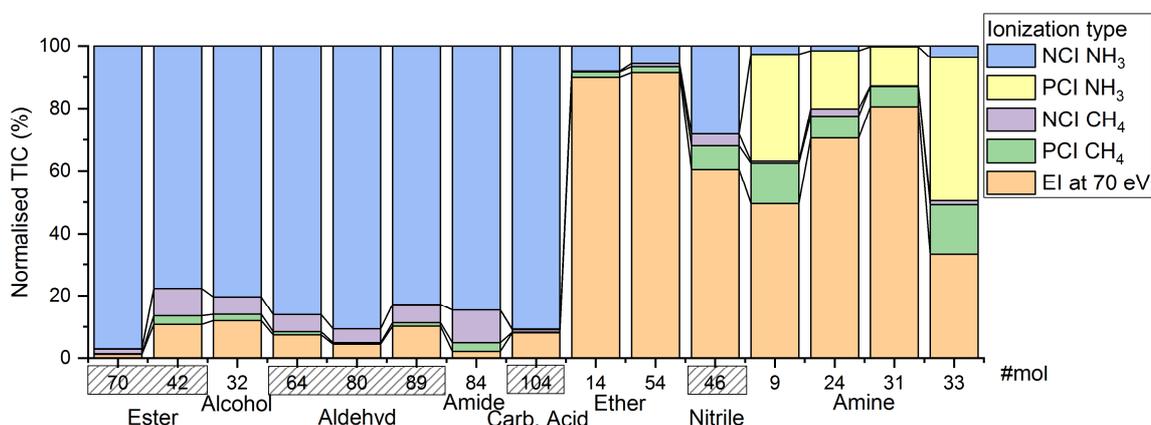


Figure 1. Percentage of TIC for each method normalized to the sum of all TIC modes (EI at 70 eV, PCI and NCI). The

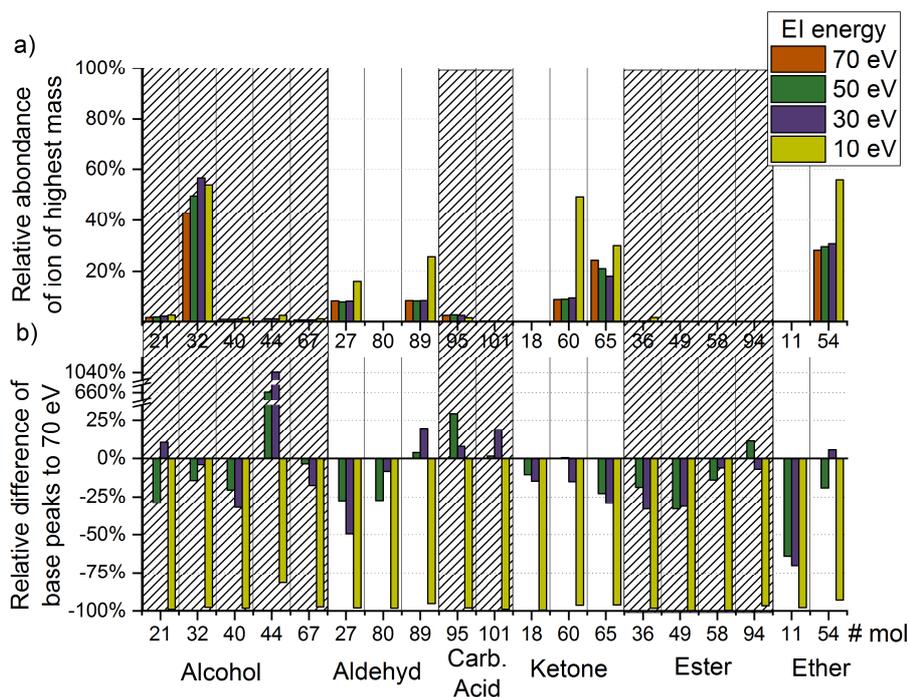
percentage of response for each method is represented for EI at 70 eV, PCI CH<sub>4</sub>, NCI CH<sub>4</sub>, PCI NH<sub>3</sub> and NCI NH<sub>3</sub>

ionizations methods (see supplementary information Table S2): orange EI at 70 eV, green PCI CH<sub>4</sub>, purple NCI CH<sub>4</sub>, yellow

PCI NH<sub>3</sub> and blue NCI NH<sub>3</sub>. Four response types can be classified into amines (with principal contribution of EI at 70 eV

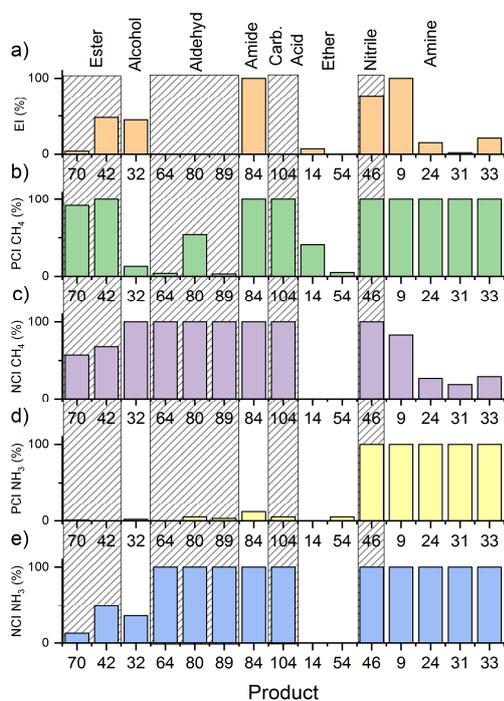
and PCI NH<sub>3</sub>), nitriles (EI at 70 eV and NCI NH<sub>3</sub>), ethers (EI at 70 eV only) and others (NCI NH<sub>3</sub> only). CH<sub>4</sub> as reagent gas

provides the lowest selectivity.



19

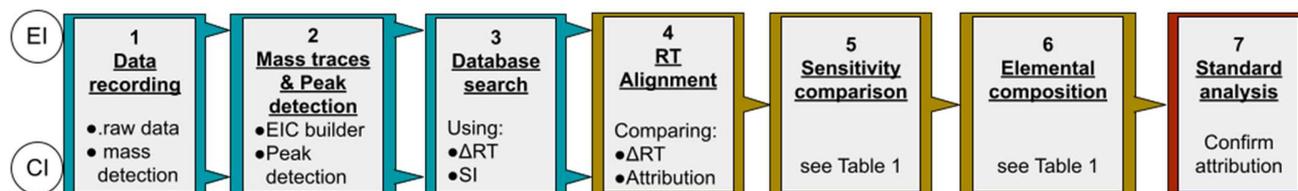
20 Figure 2. Summary of the impact of variable electron voltage (VeV) on VOCs mass pattern (molecular numbering are  
 21 extracted from supplementary information Table S1 and more detailed data are displayed in supplementary information  
 22 Table S3). Figure a) plots the intensity of ions of high mass and intensity relative to base peak and figure b) plots the relative  
 23 intensity difference of the corresponding base peak intensity with respect to 70 eV for all standards.  
 24



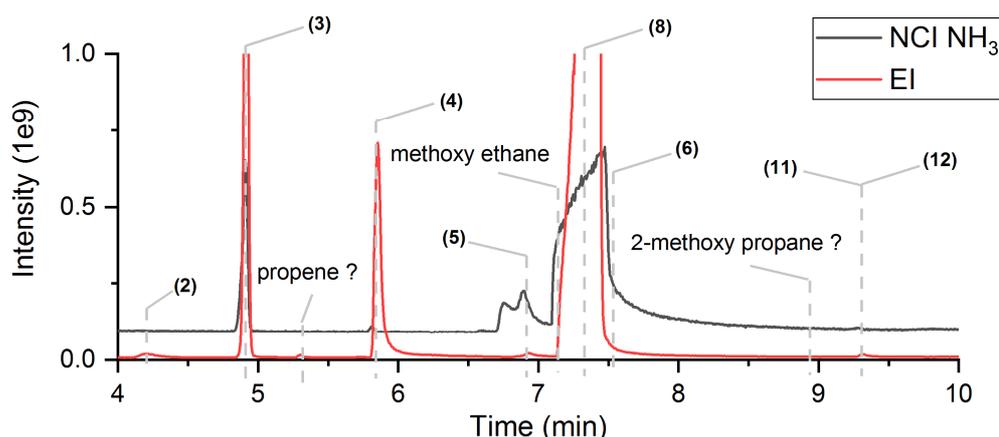
25

26 Figure 3. Relative abundance of extracted ions of high mass and intensity with respect to the ionization mode: orange

27 barcharts corresponds to electron ionization (EI), green to positive chemical ionization (PCI) using methane (PCI CH<sub>4</sub>),  
 28 purple to negative chemical ionization (NCI) using methane (NCI CH<sub>4</sub>), yellow to PCI using ammonia (PCI NH<sub>3</sub>) and blue to  
 29 NCI using ammonia (NCI NH<sub>3</sub>) (see Table S2).  
 30



31  
 32 Figure 4. FT-Orbitrap-MS Identification strategy for unknown identification after library search. The identification workflow  
 33 is the following: From the two EI and NCI analysis, step 1, mass traces are extracted building extracted ion chromatograms  
 34 (EIC) and a peak list is provided in step 2. Then each extracted spectra are compared to the database using retention time  
 35 difference ( $\Delta RT$ ) and similarity index (SI, see Eq. 1) at step 3. Because of the absence of internal standards in interstellar  
 36 analogs analysis, retention time alignment must be improved at step 4 using retention time difference and attribution  
 37 comparison. Finally, TIC sensitivity of different ionization modes can be compared, steps 5, before as the exact  
 38 fragmentation pattern analysis step 6 to propose a precise standard to analyse at step 7.



39  
 40 Figure 5. GC-MS chromatograms (TIC) from two analyses (EI at 70 eV and NCI NH<sub>3</sub>) of VOCs recovered after the warming  
 41 of an ice of pure methanol irradiated during 76 h concomitantly to the gas deposition. Identifications are labeled by following  
 42 the strategy described in Figure 4. Dioxygen (2), carbon dioxide (3), methanol (8), methyl formate (6), ethanol (12), diethyl  
 43 ether (11) and methyl acetate (22) were detected using library search and confirmed with NCI NH<sub>3</sub> mass spectra  
 44 identification. Coupling EI and NCI NH<sub>3</sub> analysis enable coelution separation (ethanol (12), diethyl ether (11)). Dimethyl  
 45 ether (4) and acetaldehyde (5), removed from the database for cross-validation, were identified using mass spectra qualitative  
 46 analysis and validated with standard identification. Peak@5.30, Peak@7.22 and Peak@9.01 qualitative analysis of HR mass  
 47 spectra are respectively consistent with C<sub>3</sub> alkanes (propene ?), methoxy ethane and 2-methoxypropane. IR spectrum is  
 48 reported in supplementary information Figure S4 and identification list and workflow are reported Table 2.

49  
 50 Table 1. Summary of the parameters used in step 5 and step 6. CI total ion current (TIC) relative to EI is the percentage of  
 51 TIC for each method normalized to the sum of EI and NCI NH<sub>3</sub> or PCI NH<sub>3</sub>.

Steps:	5a	5b	6a		6b				
	CI TIC relative to EI		ion of highest mass and intensity		NCI NH <sub>3</sub> [M-H] <sup>-</sup>		NCI NH <sub>3</sub> [M-H±2H] <sup>-</sup>		
	NCI NH <sub>3</sub>	PCI NH <sub>3</sub>	#O	#N	Rel. Ab.	DBE	Rel. Ab.	Nature	DBE
Alkanes	0 %	0 %	0	0	-	-	-	-	-
Ethers	≤ 10 %	0 %	1	0	0 %	-	-	-	-
Amines	≤ 10 %	~ 30 %	0	1	0 %	-	100 %	[M-3H] <sup>-</sup>	1.5
Nitriles	~ 33 %	0 %	0	1	≥ 90 %	2.5	≥ 8 %	[M-3H] <sup>-</sup>	3.5
Amides	≥ 90 %	0 %	1	1	-	-	-	-	-
Esters	≥ 90 %	0 %	2	0	3-100 %	1.5	-	-	-
C. Acids	≥ 90 %	0 %	2	0	≥ 90 %	1.5	-	-	-
Carbonyls	≥ 90 %	0 %	1	0	≥ 90 %	1.5	0-9 %	[M-3H] <sup>-</sup>	2.5
Alcohols	≥ 90 %	0 %	1	0	3-100 %	0.5	≥ 90 %	[M-3H] <sup>-</sup>	1.5

52 Notes: The sum of TIC for the same subset of standards is normalized to 100 and the percentage of the response for the  
53 selected ionization method (Figure 1). At step 5a three normalized TIC response types can be classified into: high, medium  
54 and low, grouping different chemical families. Step 5b allows specific identification of amines thanks to the selectivity of  
55 PCI NH<sub>3</sub>. Step 6a, the presence of O or N in multi-deprotonated anion allows differentiation of amides to esters/carboxylic  
56 acids and to carbonyls / alcohols. Then, step 6b, further deprotonate anion characteristics allows M identification with high  
57 relative abundance (Rel. Ab.) except for some esters (see supplementary information Figure S3) and amides. In particular,  
58 the use of the peak of second intensity next to the deprotonated ion (M-H±H) allows discrimination of alcohols and  
59 carbonyls (see supplementary information Table S4). In this study, only unsaturated compounds were analysed. Double bond  
60 equivalents (DBE) can thus provide information on the chemical family (see supplementary information Table S4 to Table  
61 S9).

62

63 Table 2. Ice analogs VOCs identified from the codeposition of methanol ice irradiated at Lyman α (121 nm) during 76h.

RT		Library		Workflow	
EI	NCI NH <sub>3</sub>	Identification	%EI	% NCI NH <sub>3</sub>	Identification
4.20	-	dioxygen	-	-	-
4.91	4.91	carbon dioxide	85	15	-
5.30	-	Peak@5.30	-	-	propene
5.85	5.81	Peak@5.81	98	2	dimethyl ether*
-	6.76	Peak@6.74	-	-	double peak
6.94	6.90	Peak@6.90	47	53	acetaldehyde*
7.22	-	Peak@7.22	-	-	methoxy ethane*
7.43	7.47	methanol	-	-	-
7.49	-	methyl formate	-	-	-
9.01	-	Peak@9.01	-	-	2-methoxy
-	9.27	ethanol	-	-	propane
9.31	-	diethyl ether	-	-	-
-	11.55	methyl acetate	-	-	-

64 Notes: # corresponds to the id number of the standard (see supplementary information Table S12). \*Confirmations are based  
65 on standards analysis.

66