



HAL
open science

Structural elucidation of soluble organic matter: Application to Titan's haze

Julien Maillard, Sébastien Hupin, Nathalie Carrasco, Isabelle Schmitz-Afonso,
Thomas Gautier, Carlos Afonso

► To cite this version:

Julien Maillard, Sébastien Hupin, Nathalie Carrasco, Isabelle Schmitz-Afonso, Thomas Gautier, et al.. Structural elucidation of soluble organic matter: Application to Titan's haze. *Icarus*, 2020, 340, pp.113627. 10.1016/j.icarus.2020.113627 . insu-02436390

HAL Id: insu-02436390

<https://insu.hal.science/insu-02436390>

Submitted on 8 Mar 2021

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.

Structural elucidation of soluble organic matter: application to Titan's haze

Julien MAILLARD^{1,2*}, Sébastien HUPIN², Nathalie CARRASCO¹, Isabelle SCHMITZ-AFONSO², Thomas GAUTIER¹ and Carlos AFONSO²

¹ LATMOS/IPSL, Université Versailles St Quentin, Sorbonne Université, CNRS, 11 blvd d'Alembert, F-78280 Guyancourt, France

² Université de Rouen, Laboratoire COBRA UMR 6014 & FR 3038, IRCOF, 1 Rue Tesnière, 76821 Mont St Aignan Cedex, France

***Corresponding author:**

Julien MAILLARD

Laboratoire COBRA, 1 Rue Lucien Tesnière, 76130 Mont-Saint-Aignan, France

+33 (0)2 35 52 29 19

julien.maillard@ens.uvsq.fr

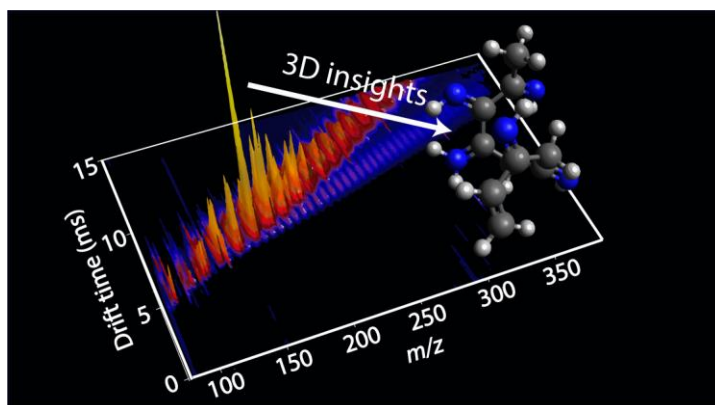
16 **Abstract**

17 The origin and evolution of organic matter in the solar system intertwines astrobiology and
18 planetary geochemistry issues. To observe the contribution of atmospheric processes in the
19 formation of complex organic matter, one of the most intensively studied objects in the outer
20 solar system is Titan, the largest moon of Saturn. Its reducing atmosphere of methane and
21 nitrogen hosts a thick, permanent, nitrogen-rich, organic haze, whose complex composition
22 remains largely unknown. Due to the measurement of species at large mass-to-charge ratio and
23 infrared information from the Cassini-Huygens mission, polyaromatic hydrocarbons (PAHs)
24 based structure for this haze has been suggested. Here, we propose a snapshot of the global
25 chemical structure based on the analysis of laboratory analogue of Titan's haze with ion
26 mobility spectrometry coupled with mass spectrometry. This robust analysis, validated with
27 other geochemical complex mixtures, such as petroleum, allows for the observation of the size
28 and three-dimensional structure of detected species. By comparison, with standards molecules,
29 we exclude several structures such as pure PolyHCN and pure polycyclic aromatic
30 hydrocarbons to be present in the principal trend of the laboratory tholins. Using theoretical
31 calculations, we propose a plausible structure consistent with our results, which is a branched
32 triazine-pyrazole. We observe that the larger the aerosols molecules are, the more they tend
33 towards a structure containing small aromatics cores linked together by short chains. We
34 suggest that the use of such an analytical approach could help advance our understanding of
35 other complex organic compounds in the Solar System such as soluble organic matter in
36 meteorites.

37 **Keywords:**

38 Titan's haze chemistry,
39 Ion mobility mass spectrometry,
40 Structural analyses,
41 Soluble organic matter,
42 Icy world,

43 **Graphical Abstract**



44

45 1. Introduction

46 Organic matter is present all over the solar system. Understanding its formation and history
47 remains a major challenge for the fields of astrobiology and planetary geochemistry. A key
48 approach to do so is to study the elemental composition and chemical structure of this organic
49 material. In this work, we propose to tackle the question of the chemical structure of complex,
50 soluble organic matters of interest for Earth field samples, meteorites, interplanetary dust
51 collected in the Earth's upper atmosphere, sample return from space missions and in situ future
52 space developments oriented towards complex organic worlds such as Titan and the ocean
53 worlds. The case study chosen here is the challenging analysis of Titan's aerosols analogues.
54 These are reduced nitrogen-rich $C_xH_yN_z$ organic samples synthesized in the laboratory from
55 processes simulating the formation of the atmospheric organic haze surrounding Titan. The
56 high degree of complexity of this material, with only carbon, nitrogen and hydrogen atoms, the
57 absence of oxygen, and its formation in cold conditions leads us to consider this model organic
58 material of interest not only for Titan, but for understanding the formation of primitive organic
59 material in the outer reduced solar system, for example comets (Jost et al., 2017) and the early
60 Earth (Trainer et al., 2006). In addition, the hydrolysis of these analogues has been shown to
61 lead to a variety of amino acids very close to the result of the Miller-Urey experiment, making
62 it a key component for astrobiology (Neish et al., 2009; Neish et al., 2010).

63 These proxies have been extensively analyzed with state-of-the-art instruments, e.g., ultrahigh
64 resolution mass spectrometers, such as FT-ICR and Orbitrap to push our understanding of this
65 organic matter (Cable et al., 2012; Danger et al., 2016; Danger et al., 2013; Somogyi et al.,
66 2005; Szopa et al., 2006; Toupance et al., 1975). Unfortunately, even with Earth best
67 instruments, revealing the structure of these analogues remains a challenge. Several analyses
68 were performed on Titan's aerosol analogues, including infrared spectroscopy (Cable et al.,
69 2014; Gautier et al., 2012; Imanaka et al., 2004), capillary electrophoresis (Cable et al., 2014)

70 and nuclear magnetic resonance (He and Smith, 2015). These previous analyses revealed the
71 presence of molecular families such as nitriles, amines and hydrocarbons. Mass spectrometry
72 analysis revealed a great deal of structural information on laboratory tholins, including a
73 polymeric trend with CH₂ and HCN as repetition pattern (Anicich et al., 2006; Bonnet et al.,
74 2013; Gautier et al., 2016; Somogyi et al., 2005; Somogyi et al., 2016; Vuitton et al., 2010).
75 Recently, mass spectrometry analyses were performed with a laser ionization desorption source
76 (Gautier et al., 2017; Maillard et al., 2018; Somogyi et al., 2012). This source allowed for the
77 comparison of both liquid and solid state of non-totally soluble tholins and showed the chemical
78 difference between these two fractions with the presence of half the amount of hydrogen in the
79 insoluble fraction (Maillard et al., 2018).

80 The main limitation of mass spectrometry analyses has been the absence of conformational and
81 isomeric information about the analyzed sample. Tandem mass spectrometry remains a solution
82 to recover structural information but is difficult to apply here due to the extreme complexity of
83 such organic samples. One way to get around this issue is to use chromatographic separation
84 (in gaseous or liquid state) upstream of the mass spectrometer to recover isomeric information.
85 The retrieved retention time is dependent of the chemical properties of the molecule and allows,
86 using standard molecules, identification of species. This technique revealed a strong potential
87 for the elucidation of several compounds such as triazole, triazine and cyanoguanidine (Gautier
88 et al., 2016). However, chromatography does not allow obtaining structural information without
89 comparison to a standard molecule and therefore remains unable to solve the structure of the
90 thousands of molecules comprising laboratory tholins.

91 Ion mobility spectrometry (IMS) is a gas phase separation method that can be coupled to mass
92 spectrometry (IMS-MS) (Mason and Schamp, 1958; Revercomb and Mason, 1975). IMS-MS
93 has been used recently for the analysis of complex mixtures such as petroleum (Castellanos et
94 al., 2014; Fernandez-Lima et al., 2009; Maleki et al., 2016). IMS is based on the separation of

95 ions in a neutral gas, according to their charge state, size and shape. Three main ion mobility
96 techniques allow the determination of collision cross section (CCS): drift tube (May et al.,
97 2014), Trapped Ions Mobility Spectrometry (TIMS) (Tose et al., 2018) and Travelling Waves
98 Ion Mobility Spectrometry (TWIMS) (Fasciotti et al., 2013; Hines et al., 2016) technology.
99 Direct measurement of the ion collision cross section (CCS) can be done with a drift tube. For
100 TIMS or TWIMS, it is possible to determine the CCS (Campuzano et al., 2012) from the
101 experimentally determined drift time (t_D) after calibration. The CCS is an intrinsic property of
102 the ions (for a particular buffer gas) that can be compared to theoretical CCS values obtained
103 from putative tridimensional structures.

104 The following study presents the advantage of IMS-MS for the study of complex organic
105 matter, in our case analogous of the Titan's hazes, to provide a snapshot of their global structure.

106 2. Methods

107 2.1. Tholins production

108 Titan's aerosol analogues, also hereafter called *tholins*, were produced with the PAMPRE
109 experiment (French acronym for *Aerosols Production in microgravity with reactive plasma*)
110 following the same procedure detailed in previous publications (Gautier et al., 2011; Szopa et
111 al., 2006). The reactor is composed of a stainless steel cylindrical reactor in which a RF-
112 Capacitively Coupled Plasma discharge is established thanks to an RF 13.56 MHz frequency
113 generator. A gas mixture containing 95% of nitrogen and 5 % of methane was injected in the
114 chamber as a continuous flow through polarized electrodes. It is then extracted by a primary
115 vacuum pump to ensure that gases are homogeneously distributed. The plasma discharge was
116 maintained at a pressure of 0.9 ± 0.1 mbar and at room temperature. A brown powder, called
117 tholins, was recovered after 1 day. The harvesting procedure was carried out under ambient air.
118 Tholins are then stored in an inert atmosphere and protected from light. Some oxidation occurs
119 during the harvest but these new oxygenated species can be sorted and are not studied during
120 the mass spectrometry analysis. It should be noticed that the pressure and temperature are lower
121 in Titan's ionosphere (respectively $\sim 10^{-5}$ - 10^{-8} mbar and 200K) than in our experiment, but the
122 ionization rate is the same (\sim ppm_v). As ion-molecule reaction rates are relatively insensitive to
123 the temperature, the lower temperature in Titan's ionosphere (200K instead of 293K in the
124 laboratory) is not an important issue in our case. The higher pressure ensures a faster kinetics
125 in the experiment without being high enough to enable molecular reactions, whereas the similar
126 ionization rate enables a realistic contribution of ions into the whole ion-neutral coupled
127 chemical network.

128

129 *2.2. Sample preparation*

130 To recover the soluble part of the sample, 4 mg of laboratory tholins were dissolved in 1 mL of methanol
131 in a vial. The vial was vigorously stirred for 3 minutes to solubilize the maximum amount of species.
132 The brown mixture was then filtered using a 0.2 µm polytetrafluoroethylene (PTFE) membrane filter on
133 a filter holder. Filtered solution was transferred in a vial and then analyzed after half dilution with a
134 50/50 water/methanol mixture to be analyzed under the same conditions than previous studies by
135 electrospray ionization. Based on electrospray mass spectrometry and elemental analysis, we
136 have evidenced the presence of less than 10% of oxygen containing species. Oxidation occurred
137 most likely during sample harvesting and storage but we cannot rule out that some oxidation
138 might occur also during ionization. The oxygen containing species have been discarded for data
139 treatment. It must be notified that around 40% of the initial quantity of laboratory tholins is
140 soluble in methanol, so the following study will only focus on a part of the sample and will not
141 be fully representative of the entire sample (Carrasco et al., 2009).

142 *2.3. Chemicals*

143 Methanol (LCMS grade) was purchased from Fisher scientific. All other chemical products (tetra-
144 alkylammonium, polyglycine and drug-like compounds) were purchased from Sigma Aldrich.

145 *2.4. Ion mobility - mass spectrometry experiments*

146 Laboratory tholins were analyzed in positive ion mode on a hybrid quadrupole-time of flight
147 mass spectrometer equipped with a Travelling Wave Ion Mobility cell (TWIMS) and
148 electrospray ion source (Synapt G2 HDMS, Waters, Manchester, UK). Although it has been
149 shown that different species ionize in negative ionization mode (Somogyi et al., 2012), it was
150 decided, in this work, to focus on the positive mode in order to present an introduction to ion
151 mobility analysis. Following ionization parameters were applied: For the source, the capillary
152 was set at 3 kV, the temperature at 100 °C, the Sampling cone at 25 V. The extraction cone was
153 also at 5 V and the desolvation temperature at 250 °C. Nitrogen cone gas flow was set at 10

154 L/Hour and nitrogen desolvation gas flow at 400 L/Hour. For Ion mobility parameters, helium
155 cell gas flow was set at 180 mL/min; Nitrogen gas flow was set at 90 mL/min, the IMS wave
156 velocity at 400 m/s. and the IMS wave height at 13 V. Trap wave velocity was set at 311 m/s
157 and Trap wave height at 6 V. Finally, transfer wave velocity was set at 191 m/s and transfer
158 wave height at 4 V. The m/z values in recorded spectra were first externally calibrated using
159 sodium formate and then internally calibrated using several well-known ions (Maillard et al.,
160 2018). Drift times were extracted for each species using DriftScope 2.8 and MassLynx 4.1.
161 Mobility dimension was calibrated using tetraalkylammonium ions (TAA) (Campuzano et al.,
162 2012) with reference values in helium. Even if the instrument is working under nitrogen, it is
163 commonly used to calibrate the CCS values in helium. Indeed, it was proven that for small
164 molecules, there is no bias in the calibration (Bleholder et al., 2015). Therefore, every
165 recovered laboratory tholins CCS is calculated in the same gas. Calibration peaks were fitted
166 with a Gaussian shape using Origin 2016. The resulting uncertainty after calibration for cross
167 sections was estimated at $\pm 0.8 \text{ \AA}^2$ by taking into account the error of the fitting (Figure S3).
168 Consequently, CCS for laboratory tholins will be stated without decimal. Table S2 shows the
169 comparison between experimental measurements and database CCS values in order to validate
170 our calibration. All reported CCS values are given following the recommendation of the ion
171 mobility community in terms of rating, calibration and results reporting (Gabelica et al., 2018).
172 Due to the resolution of the time of flight mass analyzer ($m/\Delta m$ 40 000), we choose to focus
173 on species below m/z 250 to prevent isobaric interferences and misassignments (See Figure S1-
174 S2). We have shown that at this mass range, detected species are similar in the soluble and the
175 insoluble fractions (Maillard et al., 2018).

176 2.5. Collision Cross Section calculation using theoretical model

177 From two-dimensional structures of chosen references ions, three-dimensional structures were
178 geometrically optimized (including partial charges) with Avogadro 1.1.1 using MMFF94 force

179 field, 10,000 steps, a steepest descent algorithm and a convergence of $10e^{-7}$. Conformation
180 search was performed when molecules were containing chiral center.

181 Theoretical calculation was achieved with MOBCAL (Mesleh et al., 1996; Shvartsburg and
182 Jarrold, 1996) with the trajectory method following Lennard-Jones parameters (Campuzano et
183 al., 2012) : hydrogen with atomic energy at 0.6175 meV and van der Waals distance at 2.2610
184 Å. carbon with atomic energy at 1.3266 meV and van der Waals distance at 3.0126 Å and finally
185 nitrogen with atomic energy at 1.4740 meV and van der Waals distance at 3.3473 Å. Each
186 calculated species were obtained thanks to 400,000 points (10 cycles. 1000 points in Monte
187 Carlo integrations of impact and 40 points in velocity integration.). All calculations were
188 performed considering helium as buffer gas. Helium was chosen because it is a much easier gas
189 to model than nitrogen.

190

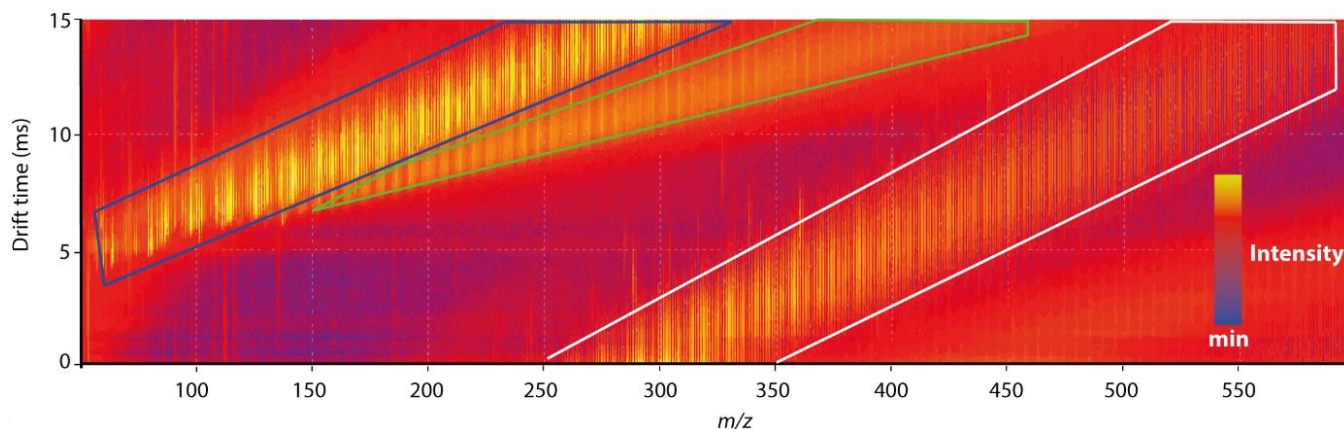
191 **3. Results**

192 *3.1. IMS-MS experiments*

193 Two pieces of observable information are recovered from the analysis of complex organic
194 matter by IMS-MS: first, the usual mass-to-charge (m/z) dimension, and then the drift time
195 dimension. The second dimension corresponds to the time taken by the ions to get across the
196 mobility cell (in practice, it includes a short period of time spent by the ions in the ion optics to
197 reach the detector). This time is characteristic of the ion size, shape and charge in the gas phase.
198 The resulting spectrum is given in Figure 1. This three-dimensional map allows a first overview
199 of the analyzed sample. The major sequence (blue area) corresponds to singly charged species.
200 In the case of laboratory tholins, a typical wave-formed structure is recovered (Bonnet et al.,
201 2013; Gautier et al., 2016; Pernot et al., 2010) with a succession of high intensity clusters
202 characteristic of known repetition patterns such as CH_2 and HCN (Somogyi et al., 2016). The
203 minor sequence below the major one on the map (green area) is composed of doubly charged
204 species naturally separated from singly charged species due to the ion mobility. These doubly
205 charged species are centered around m/z 275, which represents molecules of mass 550 u. Their
206 presence is not surprising because laboratory tholins contains a large amount of nitrogenous
207 compounds (Sciamma-O'Brien et al., 2010), which easily allow double protonation. These
208 double charged species will not be discussed in this study. A third sequence (white area) is
209 observable on Figure 1 far below the doubly charged one. This represents the larger singly
210 charged species collected from the previous scan. Indeed, these species would need a longer
211 time to be expelled from the mobility cell than the time stated for the next ion packet to be
212 released in the cell. This phenomenon is often called “wrap-around”. It could have been avoided
213 by changing parameters, such as wave velocity and wave height, but the used experimental
214 conditions were chosen to optimize the ion mobility separation in the desired drift time range.
215 In this study, we focused on the separation of ions between m/z 50 and m/z 250, as discussed in

216 the methods section. Therefore, we optimized the drift time separation for this m/z range. The
217 following results will focus on the extracted peaks in the major sequence. Their collision cross
218 section (CCS) are recovered using tetra-alkylammonium salt (TAA) ions (Campuzano et al.,
219 2012) as reference for CCS calibration in helium.

220



222 **Figure 1:** CCS vs m/z plots obtained from laboratory tholins recorded on Synapt G2 (intensity
223 in logarithm scale). Three areas are surrounded: (blue) main trend corresponding to singly
224 charged species, (green) secondary trend corresponding to doubly charged species, (white)
225 wrap around phenomenon corresponding to low mobility singly charged species coming from
226 the precedent scan.

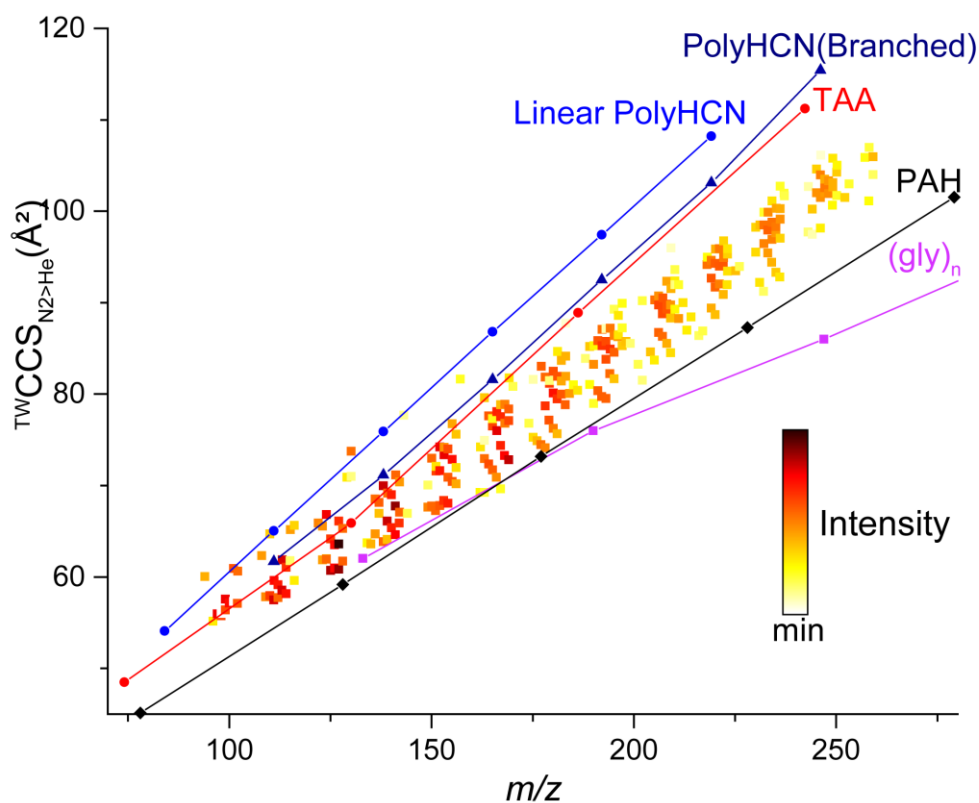
227 All recovered peaks are presented in Figure 2, which provides CCS in helium vs mass-to-charge
228 ratio for each detected peak. With this overview, we obtain the general trend for our sample.
229 To have a snapshot of the global structure of laboratory tholins sample, we added two well-
230 known families from our experimental measurements: tetra-alkylammonium ions (TAA),
231 which were used as reference CCS, and polyglycine. In addition, polycyclic aromatic
232 hydrocarbon was added from precedent measurements (PAH) (Campuzano et al., 2012; Lim et
233 al., 2018). All these families are listed in the table S2 and a comparison is given between
234 experimental measurements and database CCS in order to validate our results.

235 Furthermore, two other families are described using theoretical calculation: linear and branched
236 polyHCN (See Figure S4 for detailed formulas of plotted species).

237 PAHs represent a family of planar compounds containing only carbon and hydrogen. This
238 family is generally used in DBE vs. carbon number plot because it is the most unsaturated
239 known family (Cho et al., 2011) and was suspected to be present in Titan's aerosols and in
240 meteoritic organic matter (Waite et al., 2007). Due to the planar conformation of PAH
241 molecules, it is interesting to observe the trend of the analyzed sample in comparison to this
242 family. It can be noticed that, even if the amount of heteroatoms is not the same between
243 laboratory tholins and PAHs, the PAH line remains under all detected species of laboratory
244 tholins. This means that the planar conformation, and so PAHs, does not seem to be the
245 preferred structure of our sample at this m/z range. TAA family is composed of molecules
246 containing an ammonium core with four carbon chain arms. This group, in opposition to PAHs,
247 tends to form spherical conformations with the folding of the carbon chains. This line remains
248 above the laboratory tholins one, leading to the conclusion that this specific spherical
249 conformation is not the one preferred in our sample. Finally, polyglycine was also considered
250 as possible laboratory tholins structures. Polyglycine is a peptide, which tend to form a helical
251 structure according to our 3D optimization (See Figure S5a for a structure). The shorter Gly2
252 peptide presents a CCS close to that of tholins but larger peptides present a more compact
253 structure. Small polyglycines are not long enough to form a helical structure. This explain why
254 they fit with the tholins sample. As a short summary of the comparison of these tree families
255 with our sample, the preferential conformation present in these laboratory tholins seems to be
256 seems to be none of the three observed conformations. Two ideal PolyHCN polymers were
257 added thanks to calculated CCS as presented in the next section. The linear PolyHCN can be
258 assimilated to a straight stick shape (3D structure in figure S5b). The branched one form a non-
259 perfectly flat network (3D structure in figure S5c). These 3D structures are also given in figure

260 S5b and S5c. This family has been proposed to be the one composing Titan's aerosols. In
261 agreement with Vuitton *et al.*, 2010 (Vuitton et al., 2010), laboratory tholins are clearly not pure
262 ideal polyHCN because representative lines of this family are far above the detected sample.
263 We can then exclude this structure from our measurement. They are potentially present in small
264 quantities in the sample but not detected in these measurements.

265 In this first overview, we have recovered important information concerning the global structure
266 of laboratory tholins sample. Spherical and pure planar shapes were excluded as well as helical
267 structures.

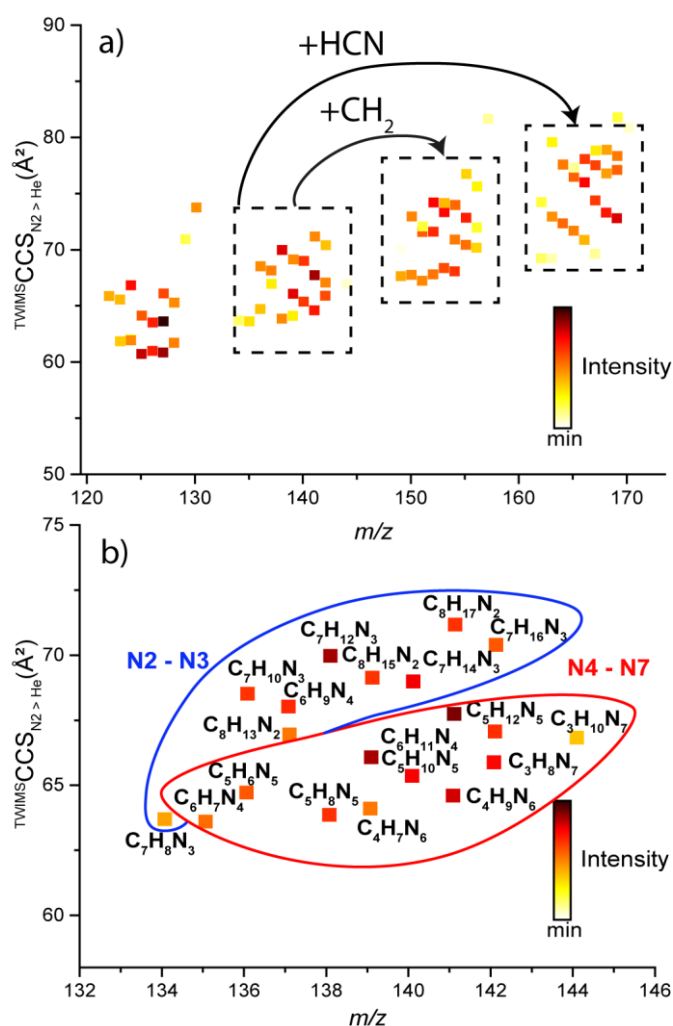


268
269 **Figure 2:** Comparison of collision cross section (Å^2) vs m/z of ions of laboratory tholins sample
270 with experimental measurements (*), polyglycin (purple) and tetraalkylammonium salts (red).
271 Taken from database (+), polyaromatics hydrocarbons (black) and calculated (o) polyHCN
272 (dashed lines)

273

274 A middle range zoom, described in Figure 3a, allows for the observation of four clusters
275 between m/z 120 and m/z 180. Ion mobility spectrometry naturally acts as a Kendrick mass
276 defect diagram (Hughey et al., 2001; Kendrick, 1963). In our case, two main repetition patterns
277 are recovered: CH_2 and HCN. This result is consistent with previous studies (Anicich et al.,
278 2006; Somogyi et al., 2005).

279 A short-range zoom is given in Figure 3b. This sub-figure focuses on a single cluster located
280 between m/z 134 and m/z 144. All molecular formulas are determined for the detected species,
281 allowing the direct observation of the collision cross section evolution according to the
282 compounds formula. Two main areas are shown according to the number of nitrogen atoms in
283 each species. The first region (in blue) is located at the top of the cluster. It contains ions with
284 a lower amount of nitrogen atoms (between N2 and N3). The other region (in red), which
285 contains ions with a higher nitrogen content (between N4 and N7) is located under the first one.
286 Therefore, the number of nitrogen atoms plays a major role in the conformation of our organic
287 matter samples and also allows the production of different isomers that are not separated with
288 the current instrumental resolution. The collision cross section decreases with an increasing
289 number of nitrogen atoms. Such evolution of polymeric species as a function of the number of
290 polar heteroelements has been observed previously (Farenc et al., 2017; Woods et al., 2004). It
291 suggests that laboratory tholins molecules have some flexibility that can yield to partial folding
292 as the number of heteroelements increase by the formation of intramolecular hydrogen bonding.
293 This overview allowed the study of the global molecular structure of the laboratory tholins
294 sample. Several typical structures (such as polyHCN, PAH, TAA and polyglycine) are excluded
295 without doubt from our sample due to the difference in the CCS values at this m/z range. After
296 working by exclusion, the next step is to use the CCS calculation tool to propose possible
297 structures of laboratory tholins.



298

299 **Figure 3:** Middle range zoom of the figure 2 for detailed comparison of collision cross section
 300 (Å^2) vs m/z of laboratory tholins ions. (a) Close zoom of the figure 2 on clusters between m/z
 301 120 and m/z 170, b) Zoom on one cluster between m/z 134 and m/z 144 with molecular formula.

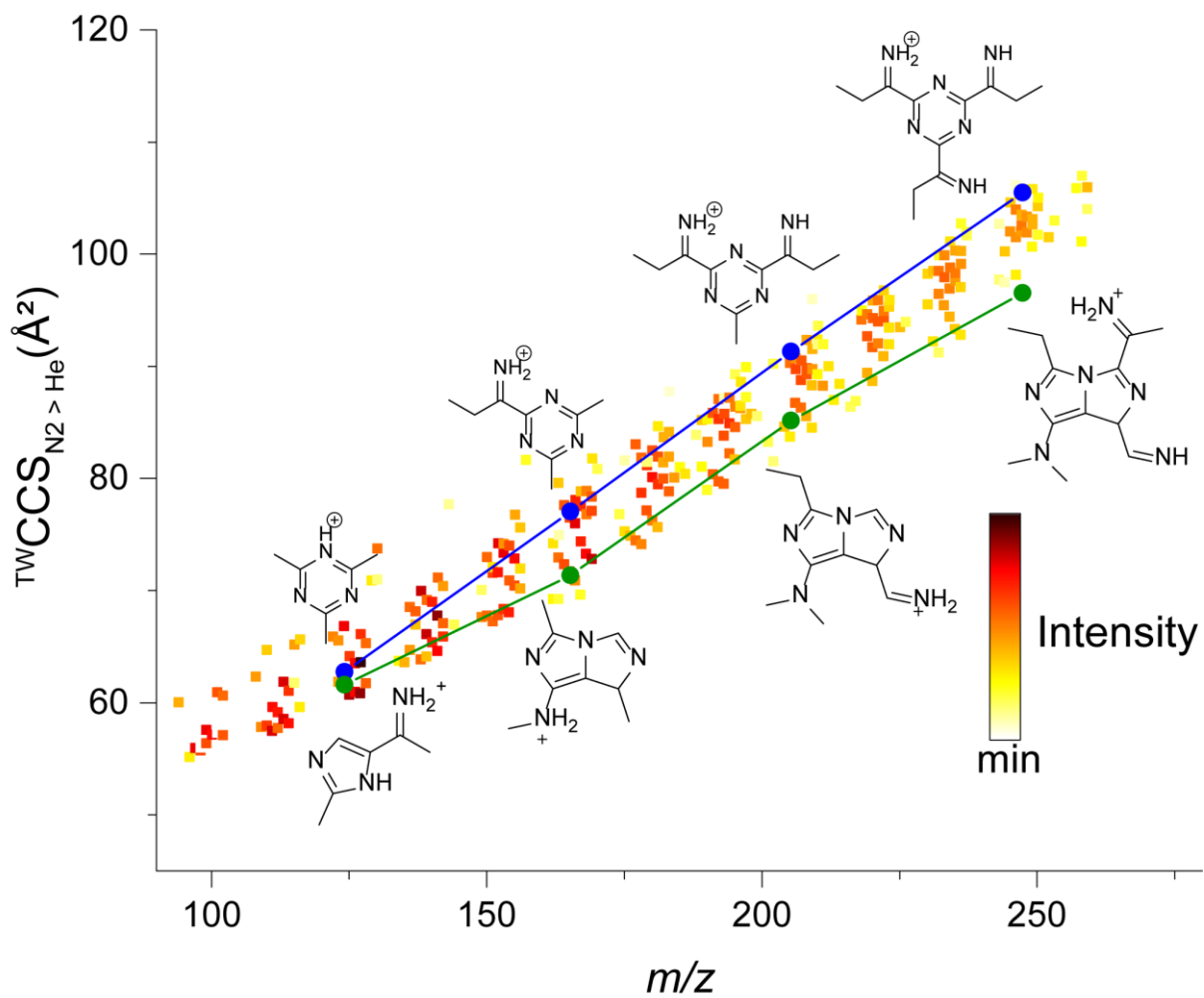
302

303 3.2. CCS calculation of potential structures

304 Due to the actual limited knowledge of such organic material molecular structures, finding a
 305 family of standard compounds that would fit the entire range of collision cross sections can be
 306 difficult. To solve this problem, CCS calculation is the easiest way to probe possible structures.

307 Theoretical CCS in helium are obtained using MOBICAL software (Mesleh et al., 1996;
 308 Shvartsburg and Jarrold, 1996) modified with new Lennard-Jones parameters (Campuzano et

309 al., 2012). The calculation method is first trained on standards to ensure its robustness (See
310 table S2 for further details). CCS values of two series of compounds are represented on Figure
311 4. The blue curve presents a family of compounds based on a triazine core substituted with
312 different numbers of branches. These branches are composed of an imine connected to a carbon
313 chain. The green curve shows a family based on pyrazole cores connected with imine groups.
314 Triazine and pyrazole cores have been chosen because their presence has been previously
315 identified in laboratory tholins (Gautier et al., 2016; Quirico et al., 2008). As observed, the CCS
316 of these two set of compounds corresponds to the CCS of tholins compounds. Thus, these two
317 families could be consistent with laboratory tholins structure, in opposition with compounds
318 presented in the previous section. It should be pointed out that several structures can match a
319 particular CCS value, the structures presented here are likely hypothesis in agreement with
320 laboratory tholins.



321

322 **Figure 4:** Comparison of CCS vs m/z with (blue) calculated CCS of triazine family (green)

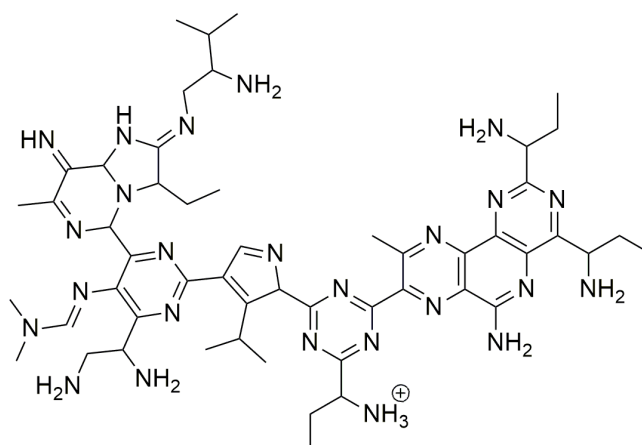
323 calculated CCS of pyrazole family.

324

325 **4. Discussion**

326 Triazine and pyrazole were identified in laboratory tholins by NMR studies (He and Smith,
327 2014) and liquid chromatography coupled to high resolution mass spectrometry (Gautier et al.,
328 2016). The first described triazine family in this study was based on these precedent works
329 (Gautier et al., 2016; He and Smith, 2014) and seems to fit well with the global trend of
330 laboratory tholins samples between PAHs and TAA families. However, the CCS values of this
331 family correspond to the highest CCS values of laboratory tholins typical ions. In opposition,
332 the second family based on a pyrazole core described in precedent works have CCS values in
333 agreement with the lowest CCS values for the laboratory tholins species. We propose, regarding
334 these results, a structure based on a Nitrogenated-PAH cores (Triazine and pyrazole) linked
335 together with short chains. The core of this set is completely conjugated. This would induce a
336 strong absorption in the UV/visible, consistent with Titan's aerosols color. A growing path
337 perspective of this family is given in Figure 5. This hypothetical structure is consistent with
338 recent works that highlighted the wide diversity of production pathways based on aromatic
339 starting cores (Gautier et al., 2017; Mahjoub et al., 2016).

340 Regarding Titan, it was proposed that the aerosols should have a PAHs structure based on
341 several observations obtained thanks to the Visual and Infrared Mapping Spectrometer (VIMS)
342 boarded on the Cassini probe (Dinelli et al., 2013; López-Puertas et al., 2013). This scenario
343 was largely approved because the aromatisation of small compounds is usually favoured due to
344 their strong stability. While our results agree with the presence of aromatic rings in the aerosol
345 structure, we observe that a pure PAH structure cannot fit the aerosol CCS. However, an
346 aromatic structure based on N-PAH cores could still be in agreement with VIMS observation
347 and would fit the structure of studied laboratory tholins sample much better.



348

349 **Figure 5:** Growing path possibility of families containing triazine and pyrazole cores. As
 350 example here, a compound with raw formula C₅₂H₇₈N₂₅ Note: This structure by itself was not
 351 detected directly in the laboratory tholins but represents a general idea of molecules that could
 352 be embedded in them.

353

354

355 **5. Conclusions**

356 In summary, this work introduces ion mobility spectrometry coupled with mass spectrometry
357 analysis for the structural study of soluble organic matter by presenting a case study on
358 analogues of the Titan's haze. This new analysis allows for an isomeric separation of the
359 sample. In addition, the collision cross section, additional structural information, is determined
360 for each ion. The collision cross sections give an overview of the main structural shape of the
361 organic matter present in the sample. Moreover, as this information is intrinsic (for a certain
362 gas) and predictable, the calculation of theoretical CCS is possible for several series of
363 compounds. Our method allows the exclusion of several structures, in the case of laboratory
364 tholins such as pure polyHCN, PAH, polyglycine and tetra-alkylammonium salts at this range
365 of mass-to-charge ratio. These are definitely not present in the major structure of laboratory
366 tholins. To conclude, we proposed two families composed of small aromatic cores linked
367 together with short chains that could be potentially be present in laboratory tholins sample. We
368 suggest using the methodology detailed here to study other material containing complex organic
369 matter, such as meteoritic soluble organic matter or Earth's kerogens. We would like to point
370 out that this first introduction provides an opportunity for further studies to complete this
371 research, including the study of negative ions as well as the insoluble fraction of tholins, both
372 proved to be an important part of the material (Maillard et al., 2018; Somogyi et al., 2012).

373

374 **Acknowledgments**

375 N.C. thanks the European Research Council (ERC) for funding via the ERC Primitive
376 Chemistry project (grant agreement No. 636829).

377 Financial support from the National Fourier transform ion cyclotron resonance network (FR
378 3624 CNRS) for conducting the research is also gratefully acknowledged.

379 This work was supported at Chimie Organique Bioorganique Réactivité Analyse (COBRA)
380 laboratory by the European Regional Development Fund (ERDF) N°31708, the Région
381 Normandie, and the Laboratoire d'Excellence (LabEx) Synthèse Organique (SynOrg) (ANR-
382 11-LABX-0029).

383 The authors wish to thank P.A. Girardi for the language proof-reading.

384

- 386 Anicich, V.G., Wilson, P.F. and McEwan, M.J. (2006) An ICR study of ion-molecules
387 reactions relevant to Titan's atmosphere: an investigation of binary hydrocarbon mixtures up
388 to 1 micron. *Journal of the American Society for Mass Spectrometry* 17, 544-561.
- 389 Bleiholder, C., Johnson, N.R., Contreras, S., Wytttenbach, T. and Bowers, M.T. (2015)
390 Molecular Structures and Ion Mobility Cross Sections: Analysis of the Effects of He and N₂
391 Buffer Gas. *Analytical chemistry* 87, 7196-7203.
- 392 Bonnet, J.-Y., Thissen, R., Frisari, M., Vuitton, V., Quirico, É., Orthous-Daunay, F.-R.,
393 Dutuit, O., Le Roy, L., Fray, N., Cottin, H., Hörst, S.M. and Yelle, R.V. (2013)
394 Compositional and structural investigation of HCN polymer through high resolution mass
395 spectrometry. *International Journal of Mass Spectrometry* 354-355, 193-203.
- 396 Cable, M.L., Hörst, S.M., He, C., Stockton, A.M., Mora, M.F., Tolbert, M.A., Smith, M.A.
397 and Willis, P.A. (2014) Identification of primary amines in Titan tholins using microchip
398 nonaqueous capillary electrophoresis. *Earth and Planetary Science Letters* 403, 99-107.
- 399 Cable, M.L., Horst, S.M., Hodyss, R., Beauchamp, P.M., Smith, M.A. and Willis, P.A. (2012)
400 Titan tholins: simulating Titan organic chemistry in the Cassini-Huygens era. *Chemical*
401 *reviews* 112, 1882-1909.
- 402 Campuzano, I., Bush, M.F., Robinson, C.V., Beaumont, C., Richardson, K., Kim, H. and
403 Kim, H.I. (2012) Structural characterization of drug-like compounds by ion mobility mass
404 spectrometry: comparison of theoretical and experimentally derived nitrogen collision cross
405 sections. *Anal Chem* 84, 1026-1033.
- 406 Carrasco, N., Schmitz-Afonso, I., Bonnet, J.Y., Quirico, E., Thissen, R., Dutuit, O., Bagag,
407 A., Laprevote, O., Buch, A., Giuliani, A., Adande, G., Ouni, F., Hadamcik, E., Szopa, C. and
408 Cernogora, G. (2009) Chemical characterization of Titan's tholins: solubility, morphology and
409 molecular structure revisited. *The journal of physical chemistry. A* 113, 11195-11203.
- 410 Castellanos, A., Benigni, P., Hernandez, D.R., DeBord, J.D., Ridgeway, M.E., Park, M.A. and
411 Fernandez-Lima, F. (2014) Fast Screening of Polycyclic Aromatic Hydrocarbons using
412 Trapped Ion Mobility Spectrometry - Mass Spectrometry. *Analytical methods : advancing*
413 *methods and applications* 6, 9328-9332.
- 414 Cho, Y., Kim, Y.H. and Kim, S. (2011) Planar limit-assisted structural interpretation of
415 saturates/aromatics/resins/asphaltenes fractionated crude oil compounds observed by Fourier
416 transform ion cyclotron resonance mass spectrometry. *Analytical chemistry* 83, 6068-6073.
- 417 Danger, G., Fresneau, A., Abou Mrad, N., de Marcellus, P., Orthous-Daunay, F.R., Duvernay,
418 F., Vuitton, V., Le Sergeant d'Hendecourt, L., Thissen, R. and Chiavassa, T. (2016) Insight
419 into the molecular composition of laboratory organic residues produced from interstellar/pre-
420 cometary ice analogues using very high resolution mass spectrometry. *Geochimica et*
421 *Cosmochimica Acta* 189, 184-196.
- 422 Danger, G., Orthous-Daunay, F.R., de Marcellus, P., Modica, P., Vuitton, V., Duvernay, F.,
423 Flandinet, L., Le Sergeant d'Hendecourt, L., Thissen, R. and Chiavassa, T. (2013)
424 Characterization of laboratory analogs of interstellar/cometary organic residues using very
425 high resolution mass spectrometry. *Geochimica et Cosmochimica Acta* 118, 184-201.
- 426 Dinelli, B.M., López-Puertas, M., Adriani, A., Moriconi, M.L., Funke, B., García-Comas, M.
427 and D'Aversa, E. (2013) An unidentified emission in Titan's upper atmosphere. *Geophysical*
428 *Research Letters* 40, 1489-1493.
- 429 Farenc, M., Witt, M., Craven, K., Barrere-Mangote, C., Afonso, C. and Giusti, P. (2017)
430 Characterization of Polyolefin Pyrolysis Species Produced Under Ambient Conditions by
431 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry and Ion Mobility-Mass
432 Spectrometry. *Journal of the American Society for Mass Spectrometry* 28, 507-514.

433 Fasciotti, M., Lalli, P.M., Klitzke, C.F., Corilo, Y.E., Pudenzi, M.A., Pereira, R.C.L., Bastos,
434 W., Daroda, R.J. and Eberlin, M.N. (2013) *Petroleomics by Traveling Wave Ion Mobility–*
435 *Mass Spectrometry Using CO₂ as a Drift Gas.* *Energy & Fuels* 27, 7277-7286.

436 Fernandez-Lima, F.A., Becker, C., McKenna, A.M., Rodgers, R.P., Marshall, A.G. and
437 Russell, D.H. (2009) *Petroleum Crude Oil Characterization by IMS-MS and FTICR MS.*
438 *Analytical chemistry* 81.

439 Gabelica, V., Afonso, C., Barran, P.E., Wyttenbach, T., Valentine, S.J., Thalassinos, K.,
440 Sobott, F., Rosu, F., Ridgeway, M.E., Richardson, K., Pagel, K., McLean, J.A., May, J.C.,
441 Kurulugama, R.T., Kim, H.I., Hann, S., Hogan, J.C.J., Groessl, M., Giles, K., Fjeldsted, J.C.,
442 Fernandez-Lima, F., Far, J., De Pauw, E., Creaser, C., Clowers, B.H., Causon, T.J., D. G.
443 Campuzano, I., Campbell, J.L., Bush, M.F., Bilbao, A., Bowers, M., T, Bleiholder, C.,
444 Benesch, J.L.P. and Shvartsburg, A.A. (2018) *Recommendations for Reporting Ion Mobility*
445 *Mass Spectrometry Measurements.*

446 Gautier, T., Carrasco, N., Buch, A., Szopa, C., Sciamma-O'Brien, E. and Cernogora, G.
447 (2011) *Nitrile gas chemistry in Titan's atmosphere.* *Icarus* 213, 625-635.

448 Gautier, T., Carrasco, N., Mahjoub, A., Vinatier, S., Giuliani, A., Szopa, C., Anderson, C.M.,
449 Correia, J.-J., Dumas, P. and Cernogora, G. (2012) *Mid- and far-infrared absorption*
450 *spectroscopy of Titan's aerosols analogues.* *Icarus* 221, 320-327.

451 Gautier, T., Schmitz-Afonso, I., Touboul, D., Szopa, C., Buch, A. and Carrasco, N. (2016)
452 *Development of HPLC-Orbitrap method for identification of N-bearing molecules in complex*
453 *organic material relevant to planetary environments.* *Icarus* 275, 259-266.

454 Gautier, T., Sebree, J.A., Li, X., Pinnick, V.T., Grubisic, A., Loeffler, M.J., Getty, S.A.,
455 Trainer, M.G. and Brinckerhoff, W.B. (2017) *Influence of trace aromatics on the chemical*
456 *growth mechanisms of Titan aerosol analogues.* *Planetary and Space Science* 140, 27-34.

457 He, C. and Smith, M.A. (2014) *A comprehensive NMR structural study of Titan aerosol*
458 *analogs: Implications for Titan's atmospheric chemistry.* *Icarus* 243, 31-38.

459 He, C. and Smith, M.A. (2015) *NMR study of the potential composition of Titan's lakes.*
460 *Planetary and Space Science* 109-110, 149-153.

461 Hines, K.M., May, J.C., McLean, J.A. and Xu, L. (2016) *Evaluation of Collision Cross*
462 *Section Calibrants for Structural Analysis of Lipids by Traveling Wave Ion Mobility-Mass*
463 *Spectrometry.* *Anal Chem* 88, 7329-7336.

464 Hughey, C.A., Hendrickson, C.L., Rodgers, R.P., Marshall, A.G. and Qian, K. (2001)
465 *Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution*
466 *Broadband Mass Spectra.* *Analytical chemistry* 73, 4676-4681.

467 Imanaka, H., Khare, B.N., Elsila, J.E., Bakes, E.L.O., McKay, C.P., Cruikshank, D.P., Sugita,
468 S., Matsui, T. and Zare, R.N. (2004) *Laboratory experiments of Titan tholin formed in cold*
469 *plasma at various pressures: implications for nitrogen-containing polycyclic aromatic*
470 *compounds in Titan haze.* *Icarus* 168, 344-366.

471 Jost, B., Pommerol, A., Poch, O., Brouet, Y., Fornasier, S., Carrasco, N., Szopa, C. and
472 Thomas, N. (2017) *Bidirectional reflectance of laboratory cometary analogues to interpret the*
473 *spectrophotometric properties of the nucleus of comet 67P/Churyumov-Gerasimenko.*
474 *Planetary and Space Science* 148, 1-11.

475 Kendrick, E. (1963) *A Mass Scale Based on CH₂= 14.0000 for High Resolution Mass*
476 *Spectrometry of Organic Compounds.* *Analytical chemistry* 35, 2146-2154.

477 Lim, D., Davidson, K.L., Son, S., Ahmed, A., Bush, M.F. and Kim, S. (2018) *Determining*
478 *Collision Cross-Sections of Aromatic Compounds in Crude Oil by Using Aromatic*
479 *Compound Mixture as Calibration Standard.* *BULLETIN OF THE KOREAN CHEMICAL*
480 *SOCIETY* 40, 122.127.

481 López-Puertas, M., Dinelli, B.M., Adriani, A., Funke, B., García-Comas, M., Moriconi, M.L.,
482 D'Aversa, E., Boersma, C. and Allamandola, L.J. (2013) Large Abundances of Polycyclic
483 Aromatic Hydrocarbons in Titan's Upper Atmosphere. *The Astrophysical Journal* 770, 132.
484 Mahjoub, A., Schwell, M., Carrasco, N., Benilan, Y., Cernogora, G., Szopa, C. and Gazeau,
485 M.-C. (2016) Characterization of aromaticity in analogues of titan's atmospheric aerosols with
486 two-step laser desorption ionization mass spectrometry. *Planetary and Space Science* 131, 1-
487 13.
488 Maillard, J., Carrasco, N., Schmitz-Afonso, I., Gautier, T. and Afonso, C. (2018) Comparison
489 of soluble and insoluble organic matter in analogues of Titan's aerosols. *Earth and Planetary*
490 *Science Letters* 495, 185-191.
491 Maleki, H., Ghassabi Kondalaji, S., Khakinejad, M. and Valentine, S.J. (2016) Structural
492 Assignments of Sulfur-Containing Compounds in Crude Oil Using Ion Mobility
493 Spectrometry-Mass Spectrometry. *Energy & Fuels* 30, 9150-9161.
494 Mason, E.A. and Schamp, H.W. (1958) Mobility of gaseous Ions in weak electric fields.
495 May, J.C., Goodwin, C.R., Lareau, N.M., Leaptrot, K.L., Morris, C.B., Kurulugama, R.T.,
496 Mordehai, A., Klein, C., Barry, W., Darland, E., Overney, G., Imatani, K., Stafford, G.C.,
497 Fjeldsted, J.C. and McLean, J.A. (2014) Conformational ordering of biomolecules in the gas
498 phase: nitrogen collision cross sections measured on a prototype high resolution drift tube ion
499 mobility-mass spectrometer. *Analytical chemistry* 86, 2107-2116.
500 Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C. and Jarrold, M.F. (1996)
501 Structural Information from Ion Mobility Measurements: Effects of the Long-Range
502 Potential. *The Journal of Physical Chemistry* 100, 16082-16086.
503 Neish, C.D., Somogyi, Á., Lunine, J.I. and Smith, M.A. (2009) Low temperature hydrolysis
504 of laboratory tholins in ammonia-water solutions: Implications for prebiotic chemistry on
505 Titan. *Icarus* 201, 412-421.
506 Neish, C.D., Somogyi, A. and Smith, M.A. (2010) Titan's Primordial Soup: Formation of
507 Amino Acids via Low-Temperature Hydrolysis of Tholins. *ASTROBIOLOGY* 10.
508 Pernot, P., Carrasco, N., Thissen, R. and Schmitz-Afonso, I. (2010) Tholinomics—Chemical
509 Analysis of Nitrogen-Rich Polymers. *Analytical chemistry* 82, 1371-1380.
510 Quirico, E., Montagnac, G., Lees, V., McMillan, P.F., Szopa, C., Cernogora, G., Rouzaud, J.-
511 N., Simon, P., Bernard, J.-M., Coll, P., Fray, N., Minard, R.D., Raulin, F., Reynard, B. and
512 Schmitt, B. (2008) New experimental constraints on the composition and structure of tholins.
513 *Icarus* 198, 218-231.
514 Revercomb, H.E. and Mason, E.A. (1975) Theory of Plasma Chromatography/Gaseous
515 Electrophoresis- A Review. *Analytical chemistry* 47.
516 Sciamma-O'Brien, E., Carrasco, N., Szopa, C., Buch, A. and Cernogora, G. (2010) Titan's
517 atmosphere: An optimal gas mixture for aerosol production? *Icarus* 209, 704-714.
518 Shvartsburg, A.A. and Jarrold, M.F. (1996) An exact hard-spheres scattering model for the
519 mobilities of polyatomic ions. *Chemical Physics Letters* 261, 86-91.
520 Somogyi, A., Oh, C.H., Smith, M.A. and Lunine, J.I. (2005) Organic environments on
521 Saturn's moon, Titan: simulating chemical reactions and analyzing products by FT-ICR and
522 ion-trap mass spectrometry. *Journal of the American Society for Mass Spectrometry* 16, 850-
523 859.
524 Somogyi, Á., Smith, M.A., Vuitton, V., Thissen, R. and Komáromi, I. (2012) Chemical
525 ionization in the atmosphere? A model study on negatively charged "exotic" ions generated
526 from Titan's tholins by ultrahigh resolution MS and MS/MS. *International Journal of Mass*
527 *Spectrometry* 316-318, 157-163.
528 Somogyi, A., Thissen, R., Orthous-Daunay, F.R. and Vuitton, V. (2016) The Role of
529 Ultrahigh Resolution Fourier Transform Mass Spectrometry (FT-MS) in Astrobiology-

530 Related Research: Analysis of Meteorites and Tholins. *International journal of molecular*
531 *sciences* 17, 439.

532 Szopa, C., Cernogora, G., Boufendi, L., Correia, J.J. and Coll, P. (2006) PAMPRE: A dusty
533 plasma experiment for Titan's tholins production and study. *Planetary and Space Science* 54,
534 394-404.

535 Tose, L.V., Benigni, P., Leyva, D., Sundberg, A., Ramirez, C.E., Ridgeway, M.E., Park,
536 M.A., Romao, W., Jaffe, R. and Fernandez-Lima, F. (2018) Coupling trapped ion mobility
537 spectrometry to mass spectrometry: trapped ion mobility spectrometry-time-of-flight mass
538 spectrometry versus trapped ion mobility spectrometry-Fourier transform ion cyclotron
539 resonance mass spectrometry. *Rapid communications in mass spectrometry : RCM* 32, 1287-
540 1295.

541 Toupance, G., Raulin, F. and Buvet, R. (1975) Formation of prebiochemical compounds in
542 models of the primitive Earth's atmosphere. *Origins of Life* 6, 83-90.

543 Trainer, M.G., Pavlov, A.A., DeWitt, H.L., Jimenez, J.L., McKay, C.P., Toon, O.B. and
544 Tolbert, M.A. (2006) Organic haze on Titan and the early Earth. *Proceedings of the National*
545 *Academy of Sciences of the United States of America* 103, 18035-18042.

546 Vuitton, V., Bonnet, J.-Y., Frisari, M., Thissen, R., Quirico, E., Dutuit, O., Schmitt, B., Le
547 Roy, L., Fray, N., Cottin, H., Sciamma-O'Brien, E., Carrasco, N. and Szopa, C. (2010) Very
548 high resolution mass spectrometry of HCN polymers and tholins. *Faraday Discussions* 147,
549 495.

550 Waite, J.H., Jr., Young, D.T., Cravens, T.E., Coates, A.J., Crary, F.J., Magee, B. and
551 Westlake, J. (2007) The process of tholin formation in Titan's upper atmosphere. *Science* 316,
552 870-875.

553 Woods, A.S., Ugarov, M., Egan, T., Koomen, J., Gillig, K.J., Fuhrer, K., Gonin, M. and
554 Schultz, J.A. (2004) Lipid/peptide/nucleotide separation with MALDI-ion mobility-TOF MS.
555 *Analytical chemistry* 76, 2187-2195.

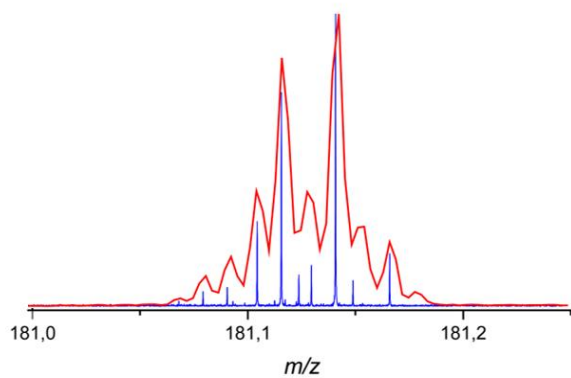
556

557

558 **Supplementary Information for**
559
560 **Structural elucidation of soluble organic matter: application to Titan's haze**

561
562 Julien MAILLARD^{†,‡,*}, Sébastien HUPIN[‡], Nathalie CARRASCO[†], Isabelle SCHMITZ-
563 AFONSO[‡], Thomas GAUTIER[†] and Carlos AFONSO[‡]
564
565 Julien MAILLARD
566 *Email:* julien.maillard@ens.uvsq.fr

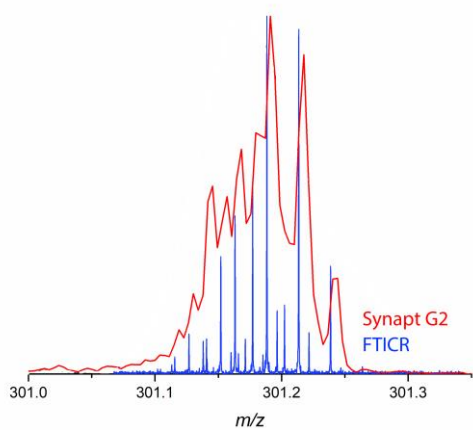
567
568 **This PDF file includes:**
569
570 Figs. S1 to S5
571 Tables S1 to S2
572



573

574

575 **Figure S1:** Comparison between Synapt G2 resolving power and FTICR at m/z 181. All species are
576 resolved with both analyser.



577

578 **Figure S2:** Comparison between Synapt G2 resolving power and FTICR at m/z 301. Above m/z 250,
579 tholins species are not resolved with the Synapt G2 analyser.

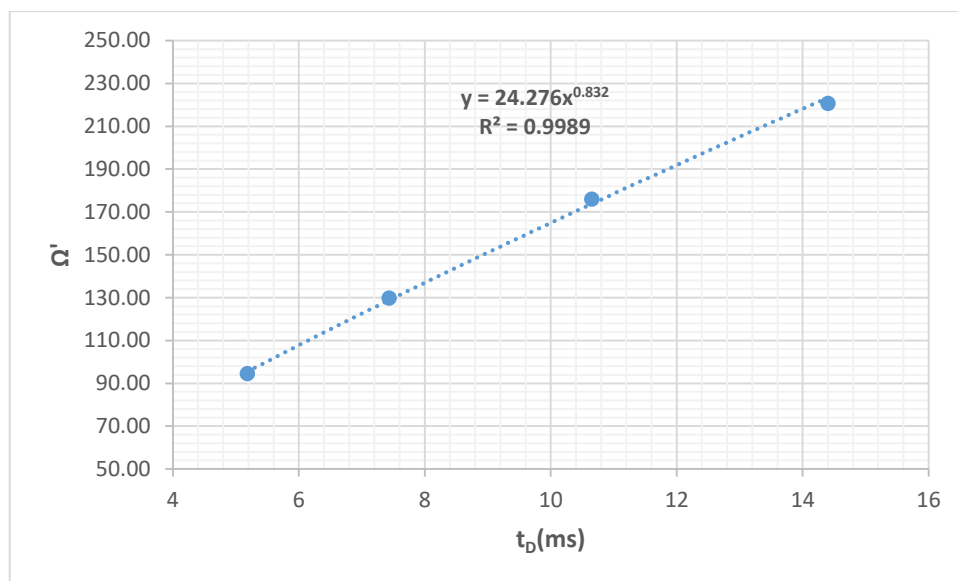
580

Table S1: TAA calibrant CCS values. Drift time was measured for each ion and added in the table.

	CCS _{He} (Å ²)	M	charge	t _D (ms)	Ω'
Methyl	48.5	74.0970	1	5.18	94.51
Ethyl	65.9	130.1596	1	7.43	129.86
Propyl	88.9	186.2222	1	10.65	175.98
Butyl	111.2	242.2848	1	14.4	220.66

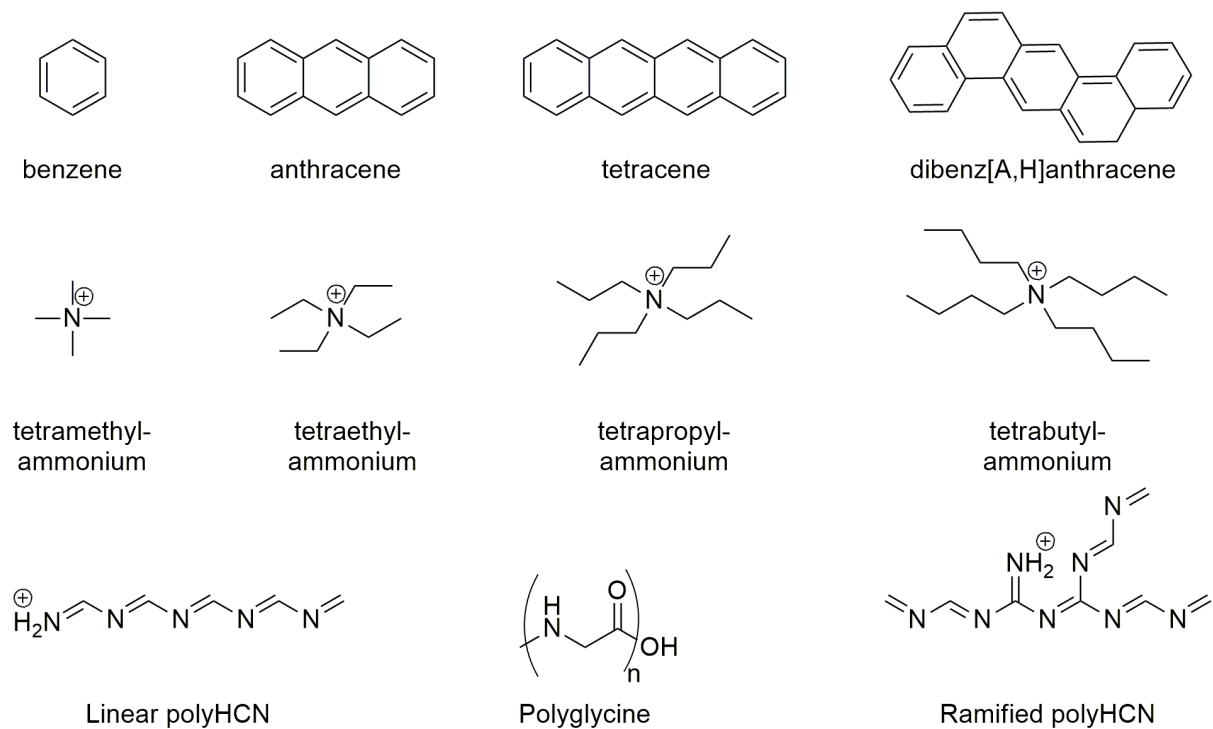
581

582



583

584 **Figure S3:** Calibration curve for the tetraalkylammonium salts in helium with a power fitting.

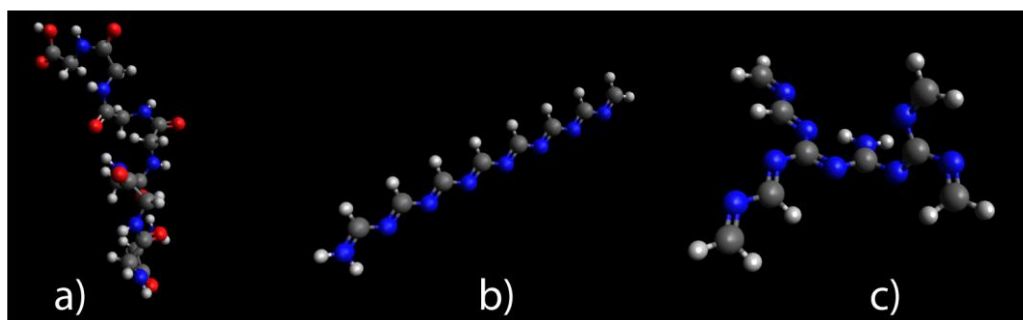


585

586 **Fig. S4.** 2D structures of compounds used for the comparison with tholins

587

588



589

590 **Fig. S5.** 3D structures of a) polyglycin, b) Linear polyHCN, c) branched polyHCN

591

Table S2: Validation of calibration parameters

Tetraalkylammonium salts (CCS calibrant)			
	Experimental CCS (Å²)		Calculated CCS (Å²)
	Literature	This work	
Methyl	49 ^a	49	48
Ethyl	66 ^a	66	66
Propyl	89 ^a	89	90
Butyl	111 ^a	111	114
Polycyclic Aromatic Hydrocarbons			
Benzene	/	/	45
Naphtalene	59 ^a	/	59
Antracene	72 ^a	/	73
dibenz[A,H]anthracene	98 ^c	/	100
Polyglycines			
Glc₂	62 ^b	64	60
Glc₃	76 ^b	76	78
Glc₄	86 ^b	87	88
Glc₅	97 ^b	/	101
Drug-like compounds			
N-Ethylaniline	63 ^a	63	62
Acetaminophen	67 ^a	69	64
Alprenolol	97 ^a	98	98
Ondansetron	106 ^a	109	108

a : Campuzano et al. Anal. Chem. 2012,

b: Wyttenbach et al. J. Am. Chem. Soc. 1998

c: Dongwan et al. Bull. Kor. Soc. 2018