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

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

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

Julien MAILLARD[1,2*], Sébastien HUPIN[2], Nathalie CARRASCO[1], Isabelle SCHMITZ-AFONSO[2], Thomas GAUTIER[1] and Carlos AFONSO[2]

[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

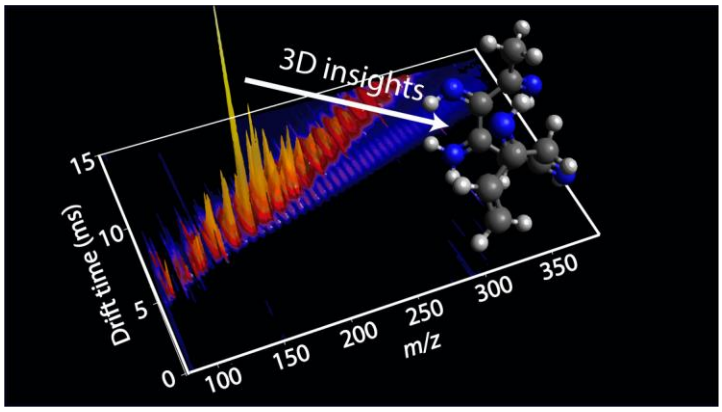
julien.maillard@ens.uvsq.fr

Abstract

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

Keywords:
Titan's haze chemistry,
Ion mobility mass spectrometry,
Structural analyses,
Soluble organic matter,
Icy world,

Graphical Abstract



1. **Introduction**

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

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

and nuclear magnetic resonance (He and Smith, 2015). These previous analyses revealed the presence of molecular families such as nitriles, amines and hydrocarbons. Mass spectrometry analysis revealed a great deal of structural information on laboratory tholins, including a polymeric trend with CH_2 and HCN as repetition pattern (Anicich et al., 2006; Bonnet et al., 2013; Gautier et al., 2016; Somogyi et al., 2005; Somogyi et al., 2016; Vuitton et al., 2010). Recently, mass spectrometry analyses were performed with a laser ionization desorption source (Gautier et al., 2017; Maillard et al., 2018; Somogyi et al., 2012). This source allowed for the comparison of both liquid and solid state of non-totally soluble tholins and showed the chemical difference between these two fractions with the presence of half the amount of hydrogen in the insoluble fraction (Maillard et al., 2018).

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

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

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

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

2. Methods

2.1. Tholins production

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

2.2. Sample preparation

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

2.3. Chemicals

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

2.4. Ion mobility - mass spectrometry experiments

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

L/Hour and nitrogen desolvation gas flow at 400 L/Hour. For Ion mobility parameters, helium cell gas flow was set at 180 mL/min; Nitrogen gas flow was set at 90 mL/min, the IMS wave velocity at 400 m/s. and the IMS wave height at 13 V. Trap wave velocity was set at 311 m/s and Trap wave height at 6 V. Finally, transfer wave velocity was set at 191 m/s and transfer wave height at 4 V. The *m/z* values in recorded spectra were first externally calibrated using sodium formate and then internally calibrated using several well-known ions (Maillard et al., 2018). Drift times were extracted for each species using DriftScope 2.8 and MassLynx 4.1. Mobility dimension was calibrated using tetraalkylammonium ions (TAA) (Campuzano et al., 2012) with reference values in helium. Even if the instrument is working under nitrogen, it is commonly used to calibrate the ccs values in helium. Indeed, it was proven that for small molecules, there is no bias in the calibration (Bleiholder et al., 2015). Therefore, every recovered laboratory tholins CCS is calculated in the same gas. Calibration peaks were fitted with a Gaussian shape using Origin 2016. The resulting uncertainty after calibration for cross sections was estimated at ±0.8 Å² by taking into account the error of the fitting (Figure S3). Consequently, CCS for laboratory tholins will be stated without decimal. Table S2 shows the comparison between experimental measurements and database CCS values in order to validate our calibration. All reported CCS values are given following the recommendation of the ion mobility community in terms of rating, calibration and results reporting (Gabelica et al., 2018).

Due to the resolution of the time of flight mass analyzer (m/Δm 40 000), we choose to focus on species below *m/z* 250 to prevent isobaric interferences and misassignments (See Figure S1-S2). We have shown that at this mass range, detected species are similar in the soluble and the insoluble fractions (Maillard et al., 2018).

2.5. Collision Cross Section calculation using theoretical model

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

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

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

3. Results

3.1. IMS-MS experiments

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

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

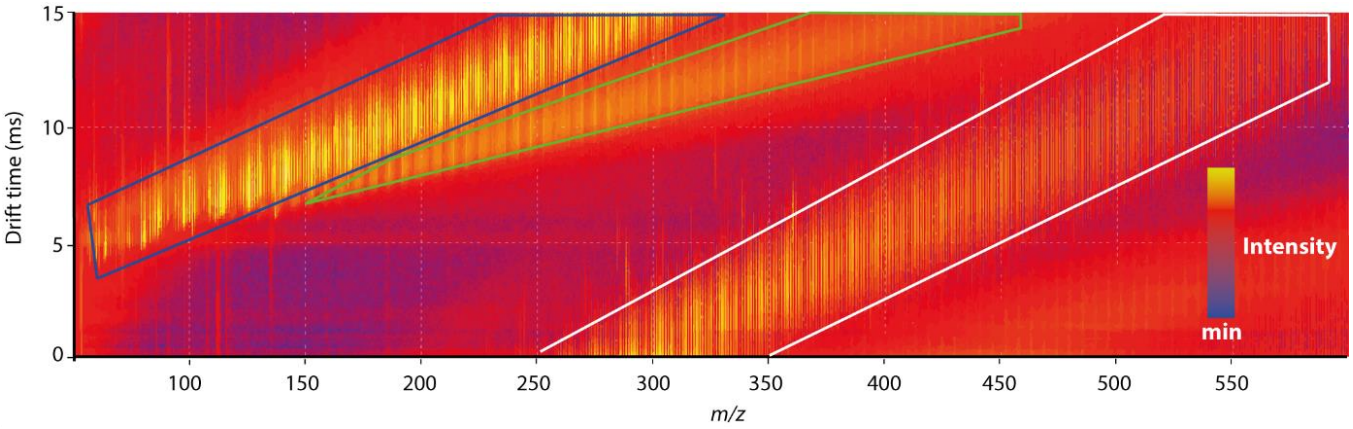


Figure 1: CCS vs *m/z* plots obtained from laboratory tholins recorded on Synapt G2 (intensity in logarithm scale). Three areas are surrounded: (blue) main trend corresponding to singly charged species, (green) secondary trend corresponding to doubly charged species, (white) wrap around phenomenon corresponding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

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

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

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

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

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

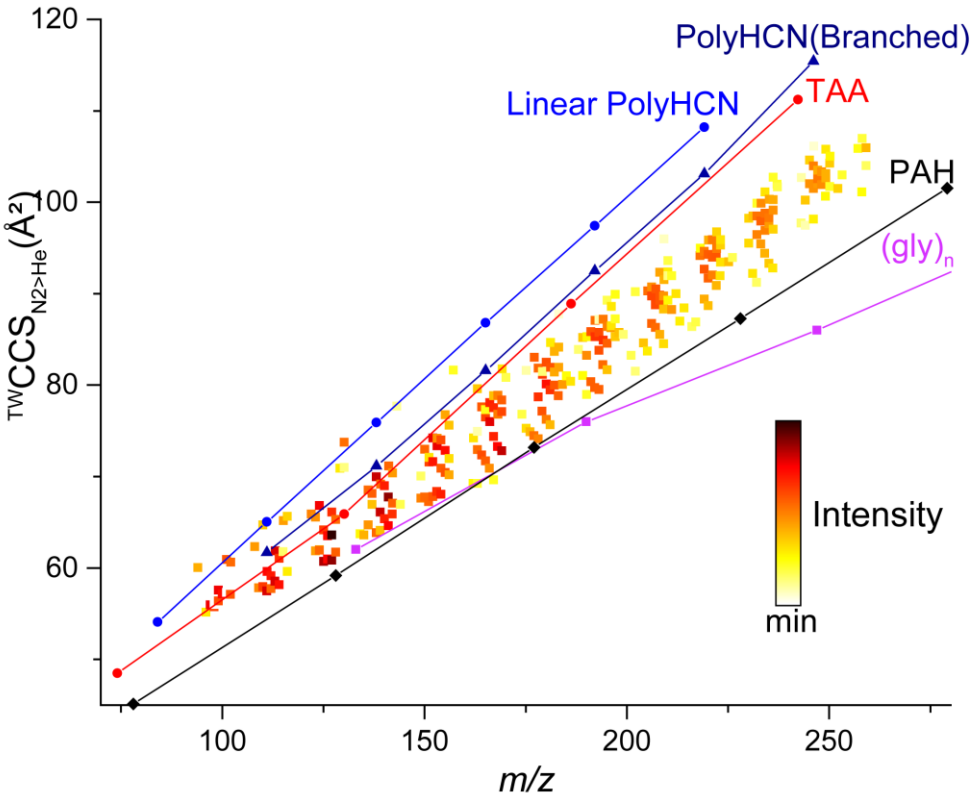
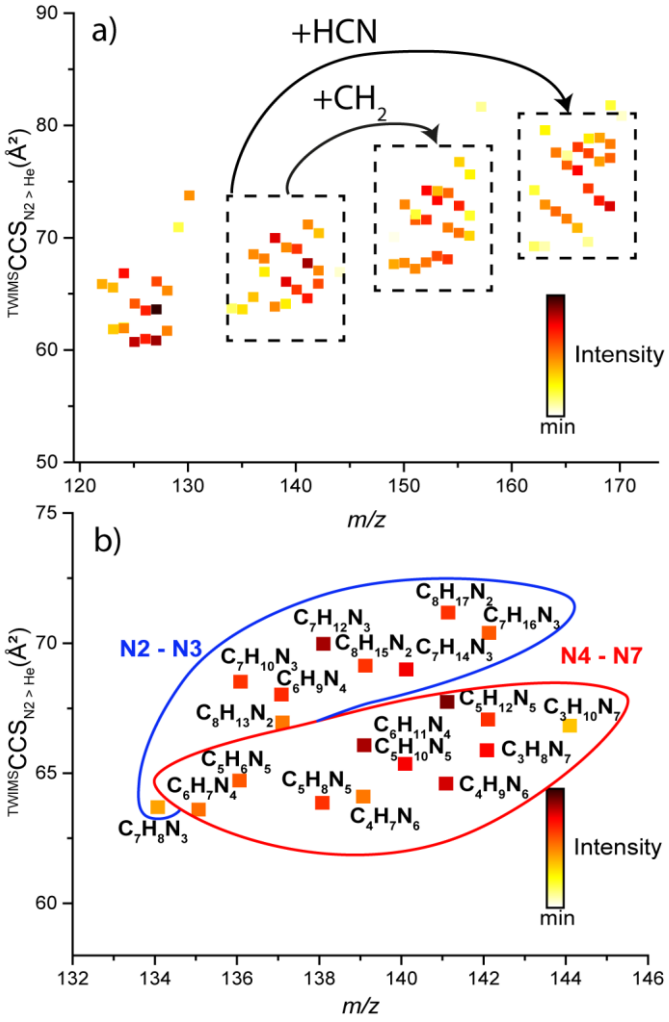


Figure 2: Comparison of collision cross section (A^2) *vs m/z* of ions of laboratory tholins sample with experimental measurements (*), polyglycin (purple) and tetraalkylammonium salts (red). Taken from database (+), polyaromatics hydrocarbons (black) and calculated (o) polyHCN (dashed lines)

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

A short-range zoom is given in Figure 3b. This sub-figure focuses on a single cluster located between *m/z* 134 and *m/z* 144. All molecular formulas are determined for the detected species, allowing the direct observation of the collision cross section evolution according to the compounds formula. Two main areas are shown according to the number of nitrogen atoms in each species. The first region (in blue) is located at the top of the cluster. It contains ions with a lower amount of nitrogen atoms (between N2 and N3). The other region (in red), which contains ions with a higher nitrogen content (between N4 and N7) is located under the first one. Therefore, the number of nitrogen atoms plays a major role in the conformation of our organic matter samples and also allows the production of different isomers that are not separated with the current instrumental resolution. The collision cross section decreases with an increasing number of nitrogen atoms. Such evolution of polymeric species as a function of the number of polar heteroelements has been observed previously (Farenc et al., 2017; Woods et al., 2004). It suggests that laboratory tholins molecules have some flexibility that can yield to partial folding as the number of heteroelements increase by the formation of intramolecular hydrogen bonding.

This overview allowed the study of the global molecular structure of the laboratory tholins sample. Several typical structures (such as polyHCN, PAH, TAA and polyglycine) are excluded without doubt from our sample due to the difference in the CCS values at this *m/z* range. After working by exclusion, the next step is to use the CCS calculation tool to propose possible structures of laboratory tholins.



298

299 **Figure 3:** Middle range zoom of the figure 2 for detailed comparison of collision cross section
300 (A²) *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.

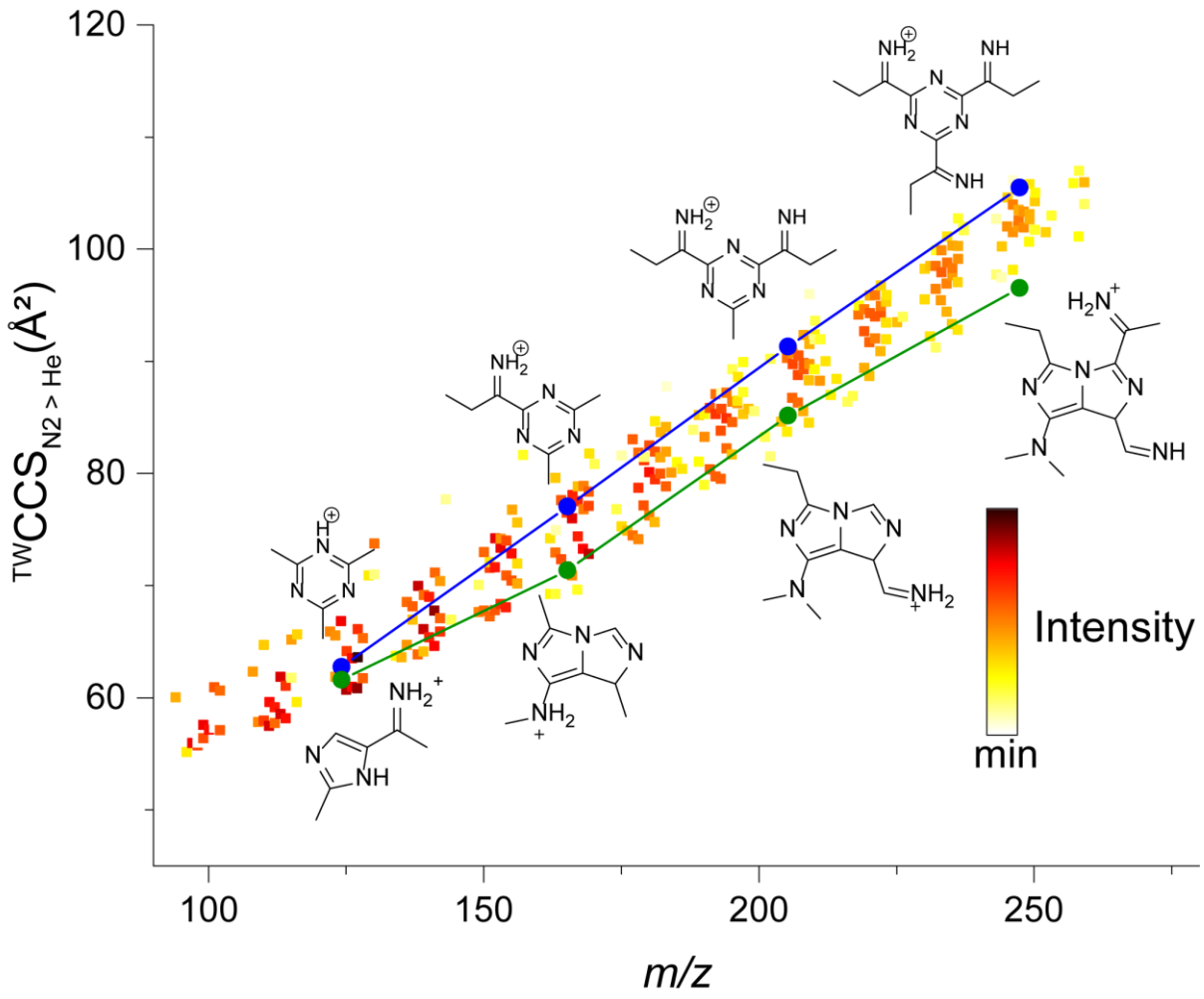
302

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 MOBCAL software (Mesleh et al., 1996;
308 Shvartsburg and Jarrold, 1996) modified with new Lennard-Jones parameters (Campuzano et

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

4. Discussion

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

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

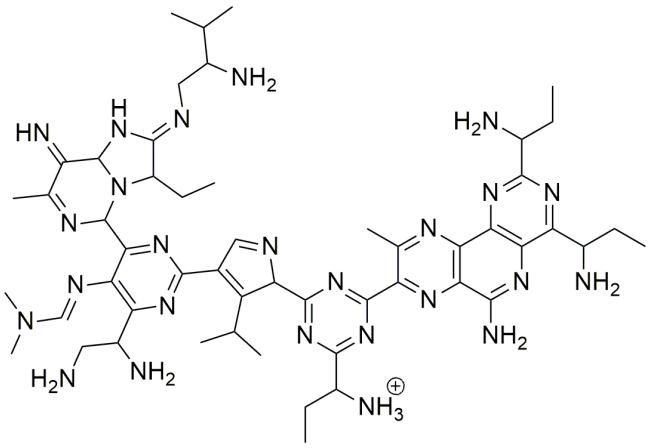


Figure 5: Growing path possibility of families containing triazine and pyrazole cores. As example here, a compound with raw formula $C_{52}H_{78}N_{25}$ Note: This structure by itself was not detected directly in the laboratory tholins but represents a general idea of molecules that could be embedded in them.

5. Conclusions

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

Acknowledgments

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

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

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

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

References

Anicich, V.G., Wilson, P.F. and McEwan, M.J. (2006) An ICR study of ion-molecules reactions relevant to Titan's atmosphere: an investigation of binary hydrocarbon mixtures up to 1 micron. Journal of the American Society for Mass Spectrometry 17, 544-561.

Bleiholder, C., Johnson, N.R., Contreras, S., Wyttenbach, T. and Bowers, M.T. (2015) Molecular Structures and Ion Mobility Cross Sections: Analysis of the Effects of He and N2 Buffer Gas. Analytical chemistry 87, 7196-7203.

Bonnet, J.-Y., Thissen, R., Frisari, M., Vuitton, V., Quirico, É., Orthous-Daunay, F.-R., Dutuit, O., Le Roy, L., Fray, N., Cottin, H., Hörst, S.M. and Yelle, R.V. (2013) Compositional and structural investigation of HCN polymer through high resolution mass spectrometry. International Journal of Mass Spectrometry 354-355, 193-203.

Cable, M.L., Hörst, S.M., He, C., Stockton, A.M., Mora, M.F., Tolbert, M.A., Smith, M.A. and Willis, P.A. (2014) Identification of primary amines in Titan tholins using microchip nonaqueous capillary electrophoresis. Earth and Planetary Science Letters 403, 99-107.

Cable, M.L., Horst, S.M., Hodyss, R., Beauchamp, P.M., Smith, M.A. and Willis, P.A. (2012) Titan tholins: simulating Titan organic chemistry in the Cassini-Huygens era. Chemical reviews 112, 1882-1909.

Campuzano, I., Bush, M.F., Robinson, C.V., Beaumont, C., Richardson, K., Kim, H. and Kim, H.I. (2012) Structural characterization of drug-like compounds by ion mobility mass spectrometry: comparison of theoretical and experimentally derived nitrogen collision cross sections. Anal Chem 84, 1026-1033.

Carrasco, N., Schmitz-Afonso, I., Bonnet, J.Y., Quirico, E., Thissen, R., Dutuit, O., Bagag, A., Laprevote, O., Buch, A., Giulani, A., Adande, G., Ouni, F., Hadamcik, E., Szopa, C. and Cernogora, G. (2009) Chemical characterization of Titan's tholins: solubility, morphology and molecular structure revisited. The journal of physical chemistry. A 113, 11195-11203.

Castellanos, A., Benigni, P., Hernandez, D.R., DeBord, J.D., Ridgeway, M.E., Park, M.A. and Fernandez-Lima, F. (2014) Fast Screening of Polycyclic Aromatic Hydrocarbons using Trapped Ion Mobility Spectrometry - Mass Spectrometry. Analytical methods : advancing methods and applications 6, 9328-9332.

Cho, Y., Kim, Y.H. and Kim, S. (2011) Planar limit-assisted structural interpretation of saturates/aromatics/resins/asphaltenes fractionated crude oil compounds observed by Fourier transform ion cyclotron resonance mass spectrometry. Analytical chemistry 83, 6068-6073.

Danger, G., Fresneau, A., Abou Mrad, N., de Marcellus, P., Orthous-Daunay, F.R., Duvernay, F., Vuitton, V., Le Sergeant d'Hendecourt, L., Thissen, R. and Chiavassa, T. (2016) Insight into the molecular composition of laboratory organic residues produced from interstellar/pre-cometary ice analogues using very high resolution mass spectrometry. Geochimica et Cosmochimica Acta 189, 184-196.

Danger, G., Orthous-Daunay, F.R., de Marcellus, P., Modica, P., Vuitton, V., Duvernay, F., Flandinet, L., Le Sergeant d'Hendecourt, L., Thissen, R. and Chiavassa, T. (2013) Characterization of laboratory analogs of interstellar/cometary organic residues using very high resolution mass spectrometry. Geochimica et Cosmochimica Acta 118, 184-201.

Dinelli, B.M., López-Puertas, M., Adriani, A., Moriconi, M.L., Funke, B., García-Comas, M. and D'Aversa, E. (2013) An unidentified emission in Titan's upper atmosphere. Geophysical Research Letters 40, 1489-1493.

Farenc, M., Witt, M., Craven, K., Barrere-Mangote, C., Afonso, C. and Giusti, P. (2017) Characterization of Polyolefin Pyrolysis Species Produced Under Ambient Conditions by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry and Ion Mobility-Mass Spectrometry. Journal of the American Society for Mass Spectrometry 28, 507-514.

Fasciotti, M., Lalli, P.M., Klitzke, C.F., Corilo, Y.E., Pudenzi, M.A., Pereira, R.C.L., Bastos, W., Daroda, R.J. and Eberlin, M.N. (2013) Petroleomics by Traveling Wave Ion Mobility–Mass Spectrometry Using CO2as a Drift Gas. Energy & Fuels 27, 7277-7286.
Fernandez-Lima, F.A., Becker, C., McKenna, A.M., Rodgers, R.P., Marshall, A.G. and Russell, D.H. (2009) Petroleum Crude Oil Characterization by IMS-MS and FTICR MS. Analytical chemistry 81.
Gabelica, V., Afonso, C., Barran, P.E., Wyttenbach, T., Valentine, S.J., Thalassinos, K., Sobott, F., Rosu, F., Ridgeway, M.E., Richardson, K., Pagel, K., McLean, J.A., May, J.C., Kurulugama, R.T., Kim, H.I., Hann, S., Hogan, J.C.J., Groessl, M., Giles, K., Fjeldsted, J.C., Fernandez-Lima, F., Far, J., De Pauw, E., Creaser, C., Clowers, B.H., Causon, T.J., D. G. Campuzano, I., Campbell, J.L., Bush, M.F., Bilbao, A., Bowers, M., T, Bleiholder, C., Benesch, J.L.P. and Shvartsburg, A.A. (2018) Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements.
Gautier, T., Carrasco, N., Buch, A., Szopa, C., Sciamma-O'Brien, E. and Cernogora, G. (2011) Nitrile gas chemistry in Titan's atmosphere. Icarus 213, 625-635.
Gautier, T., Carrasco, N., Mahjoub, A., Vinatier, S., Giuliani, A., Szopa, C., Anderson, C.M., Correia, J.-J., Dumas, P. and Cernogora, G. (2012) Mid- and far-infrared absorption spectroscopy of Titan's aerosols analogues. Icarus 221, 320-327.
Gautier, T., Schmitz-Afonso, I., Touboul, D., Szopa, C., Buch, A. and Carrasco, N. (2016) Development of HPLC-Orbitrap method for identification of N-bearing molecules in complex organic material relevant to planetary environments. Icarus 275, 259-266.
Gautier, T., Sebree, J.A., Li, X., Pinnick, V.T., Grubisic, A., Loeffler, M.J., Getty, S.A., Trainer, M.G. and Brinckerhoff, W.B. (2017) Influence of trace aromatics on the chemical growth mechanisms of Titan aerosol analogues. Planetary and Space Science 140, 27-34.
He, C. and Smith, M.A. (2014) A comprehensive NMR structural study of Titan aerosol analogs: Implications for Titan's atmospheric chemistry. Icarus 243, 31-38.
He, C. and Smith, M.A. (2015) NMR study of the potential composition of Titan′s lakes. Planetary and Space Science 109-110, 149-153.
Hines, K.M., May, J.C., McLean, J.A. and Xu, L. (2016) Evaluation of Collision Cross Section Calibrants for Structural Analysis of Lipids by Traveling Wave Ion Mobility-Mass Spectrometry. Anal Chem 88, 7329-7336.
Hughey, C.A., Hendrickson, C.L., Rodgers, R.P., Marshall, A.G. and Qian, K. (2001) Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra. Analytical chemistry 73, 4676-4681.
Imanaka, H., Khare, B.N., Elsila, J.E., Bakes, E.L.O., McKay, C.P., Cruikshank, D.P., Sugita, S., Matsui, T. and Zare, R.N. (2004) Laboratory experiments of Titan tholin formed in cold plasma at various pressures: implications for nitrogen-containing polycyclic aromatic compounds in Titan haze. Icarus 168, 344-366.
Jost, B., Pommerol, A., Poch, O., Brouet, Y., Fornasier, S., Carrasco, N., Szopa, C. and Thomas, N. (2017) Bidirectional reflectance of laboratory cometary analogues to interpret the spectrophotometric properties of the nucleus of comet 67P/Churyumov-Gerasimenko. Planetary and Space Science 148, 1-11.
Kendrick, E. (1963) A Mass Scale Based on CH2= 14.0000 for High Resolution Mass Spectrometry of Organic Compounds. Analytical chemistry 35, 2146-2154.
Lim, D., Davidson, K.L., Son, S., Ahmed, A., Bush, M.F. and Kim, S. (2018) Determining Collision Cross-Sections of Aromatic Compounds in Crude Oil by Using Aromatic Compound Mixture as Calibration Standard. BULLETIN OF THE KOREAN CHEMICAL SOCIETY 40, 122.127.

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

Related Research: Analysis of Meteorites and Tholins. International journal of molecular sciences 17, 439.
Szopa, C., Cernogora, G., Boufendi, L., Correia, J.J. and Coll, P. (2006) PAMPRE: A dusty plasma experiment for Titan's tholins production and study. Planetary and Space Science 54, 394-404.
Tose, L.V., Benigni, P., Leyva, D., Sundberg, A., Ramirez, C.E., Ridgeway, M.E., Park, M.A., Romao, W., Jaffe, R. and Fernandez-Lima, F. (2018) Coupling trapped ion mobility spectrometry to mass spectrometry: trapped ion mobility spectrometry-time-of-flight mass spectrometry versus trapped ion mobility spectrometry-Fourier transform ion cyclotron resonance mass spectrometry. Rapid communications in mass spectrometry : RCM 32, 1287-1295.
Toupance, G., Raulin, F. and Buvet, R. (1975) Formation of prebiochemical compounds in models of the primitive Earth's atmosphere. Origins of Life 6, 83-90.
Trainer, M.G., Pavlov, A.A., DeWitt, H.L., Jimenez, J.L., McKay, C.P., Toon, O.B. and Tolbert, M.A. (2006) Organic haze on Titan and the early Earth. Proceedings of the National Academy of Sciences of the United States of America 103, 18035-18042.
Vuitton, V., Bonnet, J.-Y., Frisari, M., Thissen, R., Quirico, E., Dutuit, O., Schmitt, B., Le Roy, L., Fray, N., Cottin, H., Sciamma-O'Brien, E., Carrasco, N. and Szopa, C. (2010) Very high resolution mass spectrometry of HCN polymers and tholins. Faraday Discussions 147, 495.
Waite, J.H., Jr., Young, D.T., Cravens, T.E., Coates, A.J., Crary, F.J., Magee, B. and Westlake, J. (2007) The process of tholin formation in Titan's upper atmosphere. Science 316, 870-875.
Woods, A.S., Ugarov, M., Egan, T., Koomen, J., Gillig, K.J., Fuhrer, K., Gonin, M. and Schultz, J.A. (2004) Lipid/peptide/nucleotide separation with MALDI-ion mobility-TOF MS. Analytical chemistry 76, 2187-2195.

Supplementary Information for

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

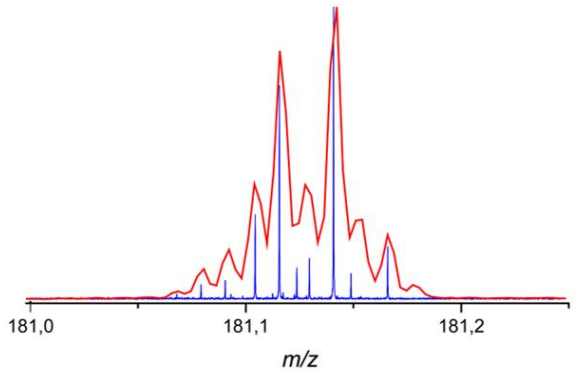
Julien MAILLARD[†,‡*], Sébastien HUPIN[‡], Nathalie CARRASCO[†], Isabelle SCHMITZ-AFONSO[‡], Thomas GAUTIER[†] and Carlos AFONSO[‡]

Julien MAILLARD

Email: julien.maillard@ens.uvsq.fr

This PDF file includes:

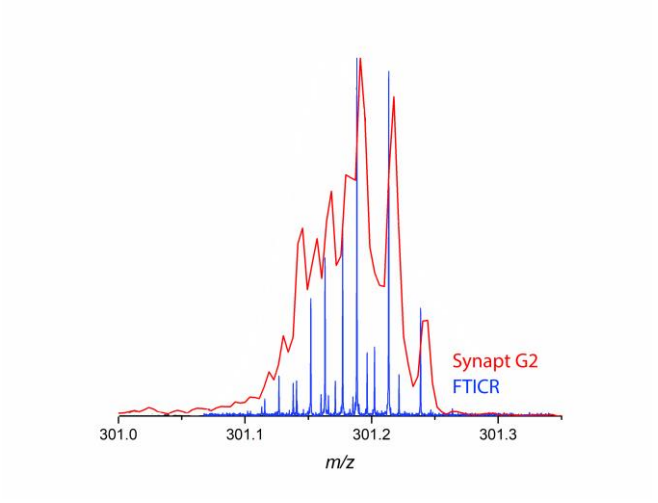
Figs. S1 to S5
Tables S1 to S2



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.

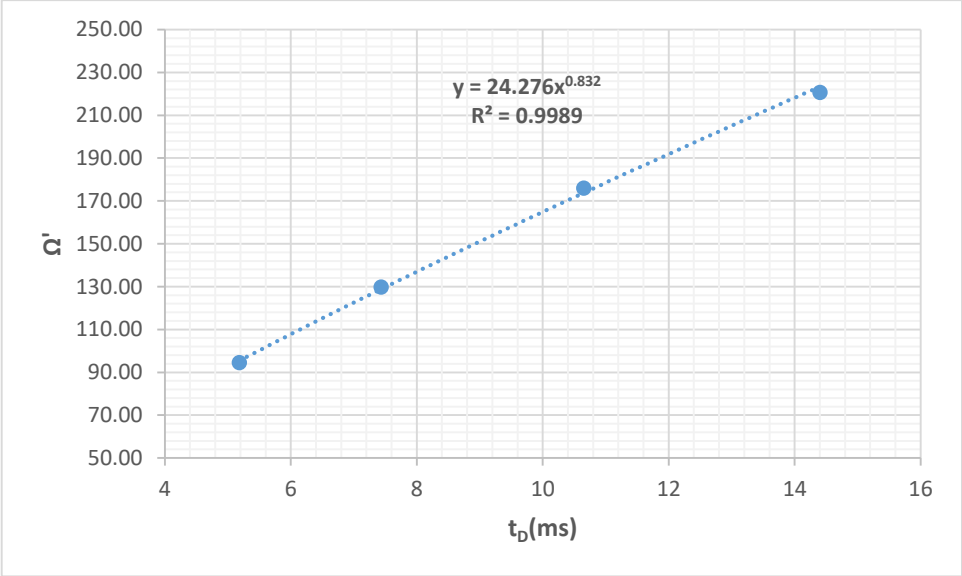
580

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

	CCS_{He} (Å²)	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 tetraalkyammonium salts in helium with a power fitting.

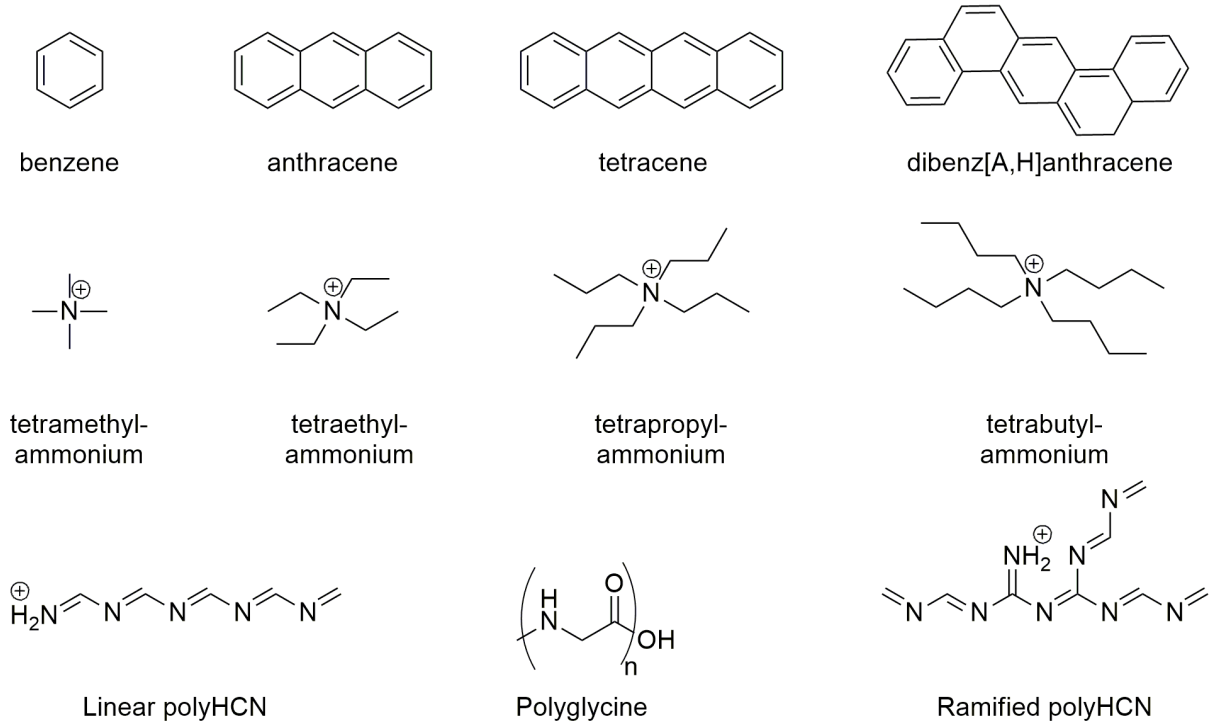


Fig. S4. 2D structures of compounds used for the comparison with tholins

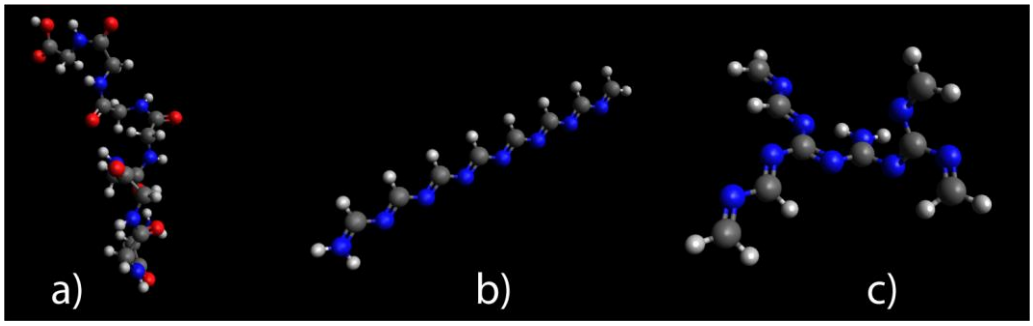


Fig. S5. 3D structures of a) polyglycin, b) Linear polyHCN, c) branched polyHCN

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			
Glc_2	62[b]	64	60
Glc_3	76[b]	76	78
Glc_4	86[b]	87	88
Glc_5	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

592