



**HAL**  
open science

## **Infrared spectroscopy quantification of functional carbon groups in kerogens and coals: A calibration procedure**

van T.H. Phan, Eric Quirico, Pierre Beck, Yann Le Brech, Lora Jovanovic, Corentin Le Guillou, Sylvain Bernard, Lydie Bonal, Nathalie Carrasco, Thomas Gautier, et al.

### ► To cite this version:

van T.H. Phan, Eric Quirico, Pierre Beck, Yann Le Brech, Lora Jovanovic, et al.. Infrared spectroscopy quantification of functional carbon groups in kerogens and coals: A calibration procedure. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* [1994-..], Elsevier, 2021, 259 (October), pp.119853. 10.1016/j.saa.2021.119853 . insu-03208201

**HAL Id: insu-03208201**

**<https://hal-insu.archives-ouvertes.fr/insu-03208201>**

Submitted on 12 Oct 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.

# Infrared spectroscopy quantification of functional carbon groups in kerogens and coals: A calibration procedure

Van T.H. Phan<sup>1</sup>, Eric Quirico<sup>1</sup>, Pierre Beck<sup>1</sup>, Yann Le Brech<sup>2</sup>, Lora Jovanovic<sup>3</sup>, Corentin Le Guillou<sup>4</sup>, Sylvain Bernard<sup>5</sup>, Lydie Bonal<sup>1</sup>, Nathalie Carrasco<sup>3</sup>, Thomas Gautier<sup>3</sup>, Jesus Raya<sup>6</sup>

<sup>1</sup>Institut de Planétologie et d'Astrophysique de Grenoble (IPAG), Université Grenoble Alpes/CNRS-INSU, UMR 5274, Grenoble F-38041,

<sup>2</sup>LRGP, CNRS, Université de Lorraine, 1 rue Grandville 54000 Nancy, France

<sup>3</sup>CNRS/INSU, LATMOS-IPSL, Guyancourt F-78280, France

<sup>4</sup>Unité Matériaux et Transformations, UMET, UMR-CNRS 8207, Université de Lille, France

<sup>5</sup>Muséum National d'Histoire Naturelle, Sorbonne Université, CNRS UMR 7590, IRD, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Paris, France

<sup>6</sup>ICS, CNRS, Université de Strasbourg, BP 296 R8 670008 Strasbourg

## Corresponding author:

Van Thi Hai Phan

Email: [thi-hai-van.phan@univ-grenoble-alpes.fr](mailto:thi-hai-van.phan@univ-grenoble-alpes.fr)

22 **ABSTRACT**

23 The determination of the abundances of the CH<sub>x</sub>, C=O and aromatic groups in chondritic  
24 Insoluble Organic Matter (IOM) and coals by Infrared (IR) spectroscopy is a challenging issue,  
25 due to insufficient knowledge on the absorption cross-sections and their sensitivity to the  
26 molecular environment. Here, we report on a calibration approach based on a <sup>13</sup>C synthetic  
27 model material, whose composition was unambiguously determined by Direct-Pulse/Magic  
28 Angle Spinning Nuclear Magnetic Resonance (DP/MAS NMR). Ratios of the cross-sections of  
29 the CH<sub>x</sub>, C=O and aromatic groups have been determined, and the method has been applied to  
30 IOM samples extracted from four chondrites as Orgueil (CI), Murchison (CM), Tagish Lake  
31 (C2-ungrouped) and EET 92042 (CR), and to a series of coals. The estimate of the aliphatic to  
32 aromatic carbon ratio ( $n_{\text{CH}_x}/n_{\text{Aro}}$ ) in IOM samples from Orgueil, Murchison and Tagish Lake  
33 chondrites is in good agreement with Single-Pulse/NMR estimates earlier published, and is  
34 lower by a factor 1.3 in the case of the CR chondrite EET 92042 (but error bars overlap). In  
35 contrast, the aliphatic to carbonyl ratio ( $n_{\text{CH}_x}/n_{\text{C=O}}$ ) is overestimated for the four chondrites.  
36 These discrepancies are likely due to the control of the absorption cross-section of the C=O and  
37 C=C bonds by the local molecular environment. Regarding coals, the use of published NMR  
38 analyses has brought to light that the integrated cross-section ratio  $A_{\text{CH}_x}/A_{\text{Aro}}$  varies with the  
39 vitrinite reflectance over an order of magnitude. Here as well, the local oxygen speciation plays  
40 a critical control on  $A_{\text{Aro}}$ , which decreases with increasing the vitrinite reflectance. We provide  
41 an analytical law that links  $A_{\text{CH}_x}/A_{\text{Aro}}$  and vitrinite reflectance, which will allow the  
42 determination of  $n_{\text{CH}_x}/n_{\text{Aro}}$  for any coal sample, provided its vitrinite reflectance is known.

43

44 **Keywords:** Infrared spectroscopy, Nuclear Magnetic Resonance (NMR), Kerogens, Coals,  
45 Insoluble Organic Matter, Chondrites, Composition

46

47

48

49

50

51

52

53

54

55

56

57

58

## 60 1. INTRODUCTION

61 Infra-Red (IR) spectroscopy is a non-destructive technique well suited to characterize  
62 complex insoluble organic compounds in meteorites and coals. It provides an estimation of  
63 organic composition based on the integrated absorbance of the main functional groups (e.g.,  
64 alkyl CH, CH<sub>2</sub> and CH<sub>3</sub>, aromatic C=C and C-H, carbonyl/carboxylic C=O, hydroxyl -OH) and  
65 tracers of the chemical and structural evolution during diagenesis and catagenesis (Li et al.,  
66 2016; Lis et al., 2005a; Painter et al., 1981; Petersen et al., 2008; Sun, 2005; Chen et al., 2012;  
67 Xueqiu et al., 2017). In addition, micro-FTIR, and more recently, Atomic force microscopy-  
68 based infrared spectroscopy (AFM-IR), has offered new perspectives to collect spatially  
69 resolved IR spectra, thereby identifying chemical heterogeneities at the submicrometric scale  
70 (Yang et al., 2017; Jakob et al., 2019; Abarghani et al., 2020; Jubb et al., 2020). These techniques  
71 have been proved valuable for the characterization of precious extraterrestrial micrometric  
72 grains, like Interplanetary Dust Particles (IDPs), Antarctic micro-meteorites and matrix grains  
73 extracted from rare primitive chondrites (Battandier et al., 2018; Beck et al., 2010; Dartois et  
74 al., 2004; Kebukawa et al., 2018; Mathurin et al., 2019; Orthous-Daunay et al., 2013; Quirico  
75 et al., 2018; Quirico et al., 2014). Although the assignment of the main vibration bands has  
76 reached a consensus, quantifying the exact proportions of functional groups based on IR signal  
77 remains difficult (Painter et al., 1981; delaRosa et al., 1992). Indeed, only a few absorption  
78 cross-sections have been published so far (Sandford et al., 1991; Dartois et al., 2004; Russell et  
79 al., 2009; Takahama et al., 2013). In addition, due to the complex structure of kerogens and  
80 coals, a given chemical group exists within different molecular environments, and is difficult to  
81 reside in various frequency shifts, inhomogeneous broadening and variability of the absorption  
82 cross-section values (Craddock et al., 2015).

83 Here, we report the determination of the absorption cross-sections of the alkyl (CH<sub>x</sub>),  
84 carbonyl (C=O) and aromatic (C=C) groups based on a H<sup>13</sup>CO synthetic kerogen analogue  
85 (named H<sup>13</sup>CO-tholins), whose composition is unambiguously determined by DP/MAS (Direct  
86 Polarization/Magic Angle Spinning) Nuclear Magnetic Resonance and XANES (X-ray  
87 Absorption at Near-Edge Spectroscopy) at the C K-edge. The use of a <sup>13</sup>C sample (e.g.  
88 synthesized H<sup>13</sup>CO-tholins) overcomes the critical bias of undetected carbons encountered in  
89 solid-state <sup>13</sup>C NMR analysis. These absorption cross-sections have been applied to a series of  
90 infrared spectra of kerogens extracted from primitive chondrites, and to a series of coals ranging  
91 a broad range of maturity. As the relative abundances of C=O, C=C and CH<sub>x</sub> groups in these  
92 samples are reported in several publications, the calibration protocol could be tested, and we  
93 further discuss the dependency of the cross-sections with the chemical structure.

94

## 95 2. MATERIALS AND METHODS

### 96 2.1 Sample selection and preparation

#### 97 2.1.1 Production of IOM analogues from HCO-tholins

98 HCO-tholins samples are polymeric disordered carbonaceous materials that were  
99 synthesized in the PAMPRE experimental reactor located at LATMOS (Guyancourt, France)  
100 (Alcouffe et al., 2010; Szopa et al., 2006). This setup consists in a radio-frequency capacitively  
101 coupled plasma (RF CCP) confined within a cylindrical cage, filled with an initial gaseous  
102 mixture made of 7% of CH<sub>4</sub>, 3% of CO and 90% of Ar gas, with a 3 sccm flux under the total  
103 pressure of 0.43 mbar. Two types of H<sup>12</sup>CO and H<sup>13</sup>CO tholins were produced from natural <sup>12</sup>C-  
104 rich and <sup>13</sup>C- rich gases (purity > 99.5%), provided by Eurisotop (Saclay, France). The power  
105 delivered by the generator was maintained within 11–12 W. Grains were formed within the  
106 plasma and kept under electrostatic levitation, and grew up until they deposited into a glass  
107 vessel surrounding the confining cage of the plasma. Experiments were run daily (8-10 h) during  
108 working hours for safety. The temperature of the electrode was 21°C when starting the  
109 discharge, and 45°C when switching it off in the evening. At total, about 30 mg of H<sup>13</sup>CO-tholins  
110 were obtained in 63 hours, and 130 mg of H<sup>12</sup>CO-tholins in 98 hours. The production yield, 0.50  
111 to 1.3 mg.hour<sup>-1</sup>, was much lower compared to tholins formed from N<sub>2</sub> : CH<sub>4</sub> mixtures  
112 (Sciamma-O'Brien et al., 2010). The H<sup>13</sup>CO-tholin was composed of a dominant yellow, fine-  
113 grained powder and a less abundant population of dark platelets (Fig. S-1). In comparison, the  
114 H<sup>12</sup>CO-tholins did not contain any dark grains. The black platelets were constituted by direct  
115 deposition on the grid, which eventually fell and mixed with the dusty tholins. Infrared  
116 spectroscopy has shown that these two populations were chemically similar (Fig. S-2). Finally,  
117 the differences in the color of the grains are probably due to the larger thickness of the platelets.

118 Analogues of kerogens and coals were obtained from the carbonization of H<sup>13</sup>CO-tholins  
119 with the elemental composition revealed in Table S-1 (Supporting Information). To select  
120 optimal conditions to get the good analogues, preliminary heating experiments were run at 200  
121 °C, 300 °C, 350 °C, 400 °C, 450 °C, 500 °C and 550 °C, with a tubular furnace maintained  
122 under secondary vacuum (~ 10<sup>-6</sup> mbar). Small grains of samples were selected and pressed  
123 between two diamond windows and put into a ceramic crucible that was placed at the middle of  
124 the tube. For each experiment, the setting temperature was maintained during 1h, and the tube  
125 was then let cool down freely for 3–6 hours, depending on the initial temperature setting.  
126 Infrared spectra of the heated samples were collected and compared to spectra of Insoluble  
127 Organic Matter (IOM) samples extracted from primitive chondrites (Fig. S-2). Finally, we select  
128 a temperature of 400°C (for extensive details on the procedure, see Supporting Information).  
129 The final carbonization of HCO-tholins was then operated with this set temperature, using a  
130 ceramic crucible filled with 5–10 mg of the powdered samples. We used an anti-static gun  
131 (Zerostat, Sigma-Aldrich) to limit sample loss due to static electricity.

132

### 133 **2.1.2 Coal and IOM samples**

134 A suite of coal samples with a mean maximum vitrinite reflectance (R<sub>o</sub>) ranging 0.33–  
135 2.8% were purchased from the Penn State University coal sample Bank  
136 (<http://www.energy.psu.edu>) (Table S-2). The IR spectra of IOMs extracted from primitive  
137 chondrites have been earlier published in Orthous-Daunay et al. (2013), including IOMs from  
138 Orgueil (CI1), Murchison (CM2), EET92042 (CR2) and Tagish Lake (C2-ung), while the IR  
139 spectra of coals were collected by using Bruker Vertex 70v spectrometer.

140

## 141 **2.2 Analytical methods**

### 142 **2.2.1 FTIR measurements**

143 The FTIR instrument used was a Bruker Hyperion 3000 infrared microscope, equipped  
144 with liquid nitrogen cooled–mercury cadmium telluride (MCT) detector. *Micro*-FTIR spectra  
145 were collected in the range of 4000 – 700  $\text{cm}^{-1}$  with a spectral resolution of 4  $\text{cm}^{-1}$ . No samples  
146 required more than 200 scans (5 min) to get a satisfying signal-to-noise ratio. All measurements  
147 were done using an environmental cell with an optical path through two ZnS windows and  
148 pumped to  $\sim 10^{-6}$  mbar (Beck et al., 2010). The sample (e.g., tholins, coal) was deposited onto  
149 a type IIa synthetic diamond window (3 mm diameter, 500  $\mu\text{m}$  thick) and crushed with a similar  
150 window to have a thin and flat sample. The diamond window holding the crushed sample was  
151 inserted in the cell, pumped, and the window was heated at 80 °C to remove efficiently adsorbed  
152 water. Background spectra were collected on a location next to the sample just before measuring  
153 the sample.

154 FTIR spectra of bulk coals and HCO-tholins samples were collected on KBr pellets.  
155 Pellets were prepared using 300 mg of a KBr powder previously heated at 300 °C under vacuum  
156 during 3 h. The sample ( $1\pm 0.1$  mg) was ground and mixed with 300 mg KBr with a MM200  
157 Retsch grinder equipped with an agate mortar filled with a single ball (30 Hz; 30 minutes).  
158 Finally, 13 mm diameter pellets were prepared using a press Atlas<sup>TM</sup> Manuelles (Eurolabo),  
159 under 400 bars and the gentle heating at 60 °C during 1h (cooling down during 30 minutes).  
160 Infrared spectra were collected under vacuum condition with a Bruker Vertex 70v spectrometer,  
161 in the range 4000 – 400  $\text{cm}^{-1}$  with 200 scans and a 4  $\text{cm}^{-1}$  spectral resolution.

162

### 163 **2.2.2 Solid-state Nuclear Magnetic Resonance**

164 The  $\text{H}^{13}\text{CO}$ - tholins material was characterized by solid NMR using an AVANCE 750  
165 MHz wide-bore spectrometer (Bruker<sup>TM</sup>) operating at MAS rate of 60 kHz designed for 1.3 mm  
166 outside diameter (OD) zirconia rotors (closed with vespel caps). About 1 – 5 mg of materials  
167 was put inside the rotor. Under those conditions (188.62 MHz and 60kHz spinning speed), any  
168 rotational sideband was thrown out of the spectral region of interest. This allows a proper  
169 deconvolution of spectra. Proton decoupling during acquisition was obtained by using SPINAL-  
170 64 at a 200 kHz  $^1\text{H}$  RF field (Fung et al., 2000). All spectra were collected within a 250kHz  
171 spectral windows (1325.3ppm) and 16324 acquisition time domains. To get undistorted  
172 lineshapes and filter out background probes signal, the quantitative Direct Polarization (DP) was  
173 performed using a speed synchronized spin-echo (Hahn, 1950; Vieira et al., 2016; Lucherelli et  
174 al., 2019), pulse sequence  $\text{P}_{90^\circ}-\tau-\text{P}_{180^\circ}-\tau$ , with the following conditions: 2.75  $\mu\text{s}$  and 5.5° s for  $\text{P}_{90^\circ}$   
175 and  $\text{P}_{180^\circ}$  pulses, and echo time being set to 33.6  $\mu\text{s}$  (2 rotation periods). Under such conditions,  
176 baseline problems and first order phasing are fixed, which is a prerequisite for quantitative  
177 analysis. T1 (spin-lattice relaxation time) were measured following inversion-recovery method  
178 (Kingsley, 1999) and led us to set up DP experiments with recycling delays equal to 30s. NMR  
179 spectra were processed with the Topspin software (phase and baseline, LB=25 applied). The  
180 interpretation of 1D  $^{13}\text{C}$  NMR spectra presents a significant difficulty due to signal overlap.

181 Therefore, the spectral decomposition of the spectra was performed with the DMFIT Software  
182 developed by Massiot et al. (2002) and previously used for biomass and biochar using Gaussian  
183 profiles (Le Brech et al., 2016). The decomposition was repeated three times for the 400 °C  
184 H<sup>13</sup>CO- tholins to estimate the uncertainty on the derived composition due to the processing.

185

### 186 2.2.3 STXM-XANES analysis

187 Synchrotron-based scanning transmission X-ray microscope (STXM) analyses were  
188 performed on heated H<sup>12</sup>CO- and H<sup>13</sup>CO-tholins to document the carbon speciation using the  
189 HERMES STXM beamline at the synchrotron SOLEIL (Saint-Aubin, France) (Belkhou et al.,  
190 2015; Swaraj et al., 2017). Samples were prepared either by cryo-ultramicrotomy (H<sup>13</sup>CO-  
191 tholins) or by depositing the powder on a Si<sub>3</sub>N<sub>4</sub> membrane (H<sup>12</sup>CO-tholin). Energy calibration  
192 was done using the well-resolved 3p Rydberg peak of gaseous CO<sub>2</sub> at 294.96 eV (Swaraj et al.,  
193 2017). XANES hypercube data (stacks) over the carbon absorption range (250 – 350 eV) were  
194 collected with a spatial resolution of 35 nm at energy increments of 0.1 eV over the 283 – 291.5  
195 eV region, with a dwell time of less than 1 ms per pixel to prevent irradiation damage (Wang et  
196 al., 2009). Stack alignments and extraction of XANES spectra were done using the aXis2000  
197 software (ver2.1n). Normalization of data and quantification of the concentrations of functional  
198 groups were done using the QUANTORXS freeware  
199 (<https://github.com/CorentinLG/QuantORXS/>) described in Le Guillou et al. (2018). Note that  
200 although this quantification procedure provides consistent results for aromatic, ketone\phenol  
201 and carboxylic\ester groups, it remains difficult to estimate the concentration of aliphatic groups  
202 precisely.

203

### 204 2.2.4 Raman setup

205 Raman measurements were collected with a HORIBA JOBIN-YVON LaRam  
206 spectrometer, equipped with a 600 gr.mm<sup>-1</sup> groove and operating at 532 nm (INSU national  
207 Facility at ENS-Lyon, France). The laser beam was focused through a ×100 objective. The  
208 power onto the sample was 380 μW, and acquisition time was 5×20s. The configuration used  
209 yields a lateral resolution of ≈ 1 μm.

210

### 211 2.2.5 Analysis of infrared spectra

212 For each spectrum, a baseline correction was done using spline functions, and the peak  
213 absorbance of the C=C band at 1540 cm<sup>-1</sup> was set to 1. The C=O and C=C bands in H<sup>13</sup>CO-  
214 tholin spectra at 1656 and 1540 cm<sup>-1</sup>, respectively, were fitted with Gaussian profiles (using the  
215 multi-fit peak functions of the Igor 8 software, Wavemetrics) (Fig.1).

216 The integrated cross-section is defined by:

217 
$$A = \int_{\nu_1}^{\nu_2} \sigma d\nu = \int_{\nu_1}^{\nu_2} \ln(10) \frac{a(\nu) \times S}{n} d\nu \text{ (cm/group) (Equation 1)}$$



218 where  $a(\nu) = -\log(\frac{I}{I_0})$  is the absorbance, with  $I$  and  $I_0$  are the intensity of the infrared beam  
219 after and before the sample, respectively;  $n$  is the number of molecular oscillators and  $\sigma$  is the  
220 absorption cross-section ( $\text{cm}^2$ ), with  $\nu$  the wavenumber,  $\nu_1$  and  $\nu_2$  the edges of the spectral range  
221 of the spectral component, and  $S$  is the probed area ( $\text{cm}^2$ ).

222 Finally, the ratio of the absorption cross-section of two chemical groups  $X_1$  and  $X_2$  is:

223 
$$\frac{A_{X1}}{A_{X2}} = \frac{a_{X1}}{a_{X2}} \times \frac{n_{X2}}{n_{X1}} \quad (\text{Equation 2})$$

224 where  $a_X$  refers to the integrated absorbance of a chemical group  $X$  as aliphatic  $\text{CH}_x$ ,  $\text{C}=\text{O}$   
225 or aromatic species (Aro).  $a_{\text{CH}_x}$  was calculated from the integration of the alkyl band between  
226 2750 and 3000  $\text{cm}^{-1}$ , while  $a_{\text{CO}}$  and  $a_{\text{Aro}}$  were obtained from a Gaussian fit of the bands at 1656  
227 and 1540  $\text{cm}^{-1}$ , respectively. The integrated cross-section ratios of the  $\text{CH}_x$ ,  $\text{C}=\text{O}$  and Aro  
228 molecular oscillators,  $A_{\text{CH}_x}/A_{\text{Aro}}$ ,  $A_{\text{CH}_x}/A_{\text{C}=\text{O}}$  or  $A_{\text{Aro}}/A_{\text{C}=\text{O}}$  were finally calculated  
229 with Equation 2, using estimates from NMR measurements on the 400 °C  $\text{H}^{13}\text{CO}$ -tholin. No  
230 estimation of individual  $A_X$  cross-sections was made, because NMR measurements provided  
231 ratios of chemical groups abundances.

232 Strictly speaking, IR spectroscopy probes normal modes and not bonds. In the case of the  
233  $\text{CH}_x$  groups, the mass difference between H and C atoms is large and decouples their respective  
234 motions. Thus, the corresponding bands are controlled by hydrogen vibration around a single  
235 carbon atom. For the terminating carbonyl group, environmental effects from neighboring  
236 chemical groups are present, leading to a moderate frequency shift and variations of the  
237 absorption coefficient by up to a factor of 2. But for both  $\text{CH}_x$  and  $\text{C}=\text{O}$ , the number of bonds  
238 or molecular groups is equal to the number of carbon atoms and a straightforward comparison  
239 with  $^{13}\text{C}$ -NMR is possible. Things get harder in the case of aromatic species. A benzene single  
240 ring ( $\text{C}_6\text{H}_6$ ), in the gas or as a molecular crystal has a vibration mode at  $\sim 1500 \text{ cm}^{-1}$  that involves  
241 all carbon atoms (6). Small polycyclic hydrocarbons (1-4 rings) display a sharp narrow band in  
242 the range 1700-1500  $\text{cm}^{-1}$ , but the number of  $\text{C}\cdots\text{C}$  bonds does not equal the number of C atoms.  
243 For instance, naphthalene units have 10 C atoms and 11  $\text{C}\cdots\text{C}$  bonds, anthracene and  
244 phenanthrene have 14 C atoms and 16  $\text{C}\cdots\text{C}$  bonds, and pyrene 16 C atoms and 19  $\text{C}\cdots\text{C}$  bonds.  
245 In the case of larger polycyclic aromatics, theoretical calculations point to a broad feature  
246 peaking around 1600  $\text{cm}^{-1}$ , which is due to the  $E_{2g}$  normal mode that involves all carbon atoms  
247 and phonon propagation (Mapelli et al., 1999). Immature kerogens and coals contain small (1-  
248 4 rings) polycyclic aromatics inserted in a macromolecular network and display a similar broad  
249 band at 1600  $\text{cm}^{-1}$ . Altogether, the infrared absorption at 1600  $\text{cm}^{-1}$  results from the motion of  
250 all carbon atoms, but in contrast no clear connection can be established with a number of  $\text{C}=\text{C}$   
251 bonds, nor a number of rings. In this respect, the terms “aromatic ring stretch” or “ $\text{C}=\text{C}$ ” used  
252 in literature are somewhat misleading, and we have used the term “Aro”. As such,  $A_{\text{Aro}}$  is  
253 controlled by the number of carbon atoms, allowing a direct comparison with  $^{13}\text{C}$ -NMR. In the  
254 details, we expect significant variations of  $A_{\text{Aro}}$  with the size and local environment of the  
255 polyaromatic specie, for instance  $-\text{OH}$  groups (see discussion below).

256 Last, the Artificial Bees algorithm was used from the ESA Package PYGMO  
257 (<https://esa.github.io/pygmo/>) and tested two fit models: one based on 7 Gaussian and one on 7



258 pseudo-Voigt components. Each group comprised 50 simulations and the fit quality was  
259 estimated with a  $\chi^2$  function. A 20000 bees' population over 50 generations was used. For each  
260 of the 200 numerical simulations to get the good fit in term of  $\chi^2$  value.

261

## 262 3 RESULTS

### 263 3.1 Solid-state $^{13}\text{C}$ DP/MAS NMR analysis

264 Solid-state  $^{13}\text{C}$  DP/MAS NMR spectra of the pristine and heated  $\text{H}^{13}\text{CO}$ -tholins are  
265 presented in Fig. 2a and the band assignment is given in Table 1, following the fitting results in  
266 Fig. 2b and 1c for the initial and 400 °C  $\text{H}^{13}\text{CO}$  tholins, respectively. We observe two intense  
267 features at 100–160 ppm and 0 – 100 ppm in the spectra of both samples, which are assigned to  
268  $\text{sp}^2$  and  $\text{sp}^3$  carbon atoms, respectively. The range 0–45 ppm usually points to C atoms in alkyl  
269 chains, while the range of 45-90 ppm points to oxygenated aliphatic carbons, e.g. alcoholic,  
270 carboxylic or ether groups (Pretsch et al., 2000). It is, however, challenging to separate the  
271 respective contribution of these two groups of carbon atoms accurately.  $\text{sp}^2$  carbons in the 100-  
272 160 ppm range can be assigned to olefinic chains (conjugated and unconjugated alkenes) and  
273 aromatic or polyaromatic species. We cannot distinguish these two types of speciation. The  
274 broad and weak signal which peaks at 170 ppm points to C=O groups in esters or carboxylic  
275 groups, and that at 210 ppm indicates to C=O in ketones or aldehydes. The effect of heating  
276 results in a steady increase of unsaturated versus saturated carbons, and a dramatic drop of labile  
277 oxygenated groups (Fig. 2a). This is in full agreement with the carbonization process that  
278 promotes hydrogen loss, carbonyl loss and aromatization.

279 The relative abundances  $F_x$ , of the different carbon groups, was derived through a  
280 decomposition procedure using a fit model based on Gaussian components (Fig. b-c; Table 1)  
281 (Massiot et al., 2002, Le Brech et al., 2016). The unheated  $\text{H}^{13}\text{CO}$ - tholins appears to be mainly  
282 composed of  $-\text{CH}_x$  ( $x=1$  to 3) groups ( $F_{\text{ali}} = \sim 61\%$ ), olefinic carbons ( $F_{\text{aro}} = \sim 32.5\%$ ) and a  
283 lower content of C=O groups ( $F_{\text{CO}} = 5.1\%$ ). The  $^{13}\text{C}$  NMR spectrum of the 400 °C heated  
284  $\text{H}^{13}\text{CO}$ -tholins is dominated by aromatic carbons ( $F_{\text{aro}} = \sim 64\%$ ) and contains less aliphatic-CH  
285 groups ( $F_{\text{ali}} = \sim 32\%$ ) and C=O groups ( $F_{\text{CO}} = 4.1\%$ ) (Table 1).

286

### 287 3.2 Raman Analysis

288 The Raman spectrum of the unheated  $\text{H}^{13}\text{CO}$ -tholin is dominated by a strong  
289 fluorescence background, with no detectable vibrational bands. The spectra of the heated  
290  $\text{H}^{13}\text{CO}$ -tholin display the first-order carbon bands superimposed onto a high fluorescence  
291 background. The spectral contrast (i.e. the intensity of the G and D bands with respect to the  
292 fluorescence intensity) is low, leading to poor quality of the reduced spectra after subtraction of  
293 a linear baseline (Fig. S-3). The use of a Lorentzian-Breit-Wigner-Fano fit model was  
294 complicated, due to the poor spectral quality at the edge of the G and D bands. Despite these  
295 troubles, these data point a G band peak position at  $\sim 1537\text{ cm}^{-1}$  and a FWHM of  $\sim 80\text{ cm}^{-1}$ . For  
296 comparison with published data, the peak position of a  $^{12}\text{C}$ -bearing sample can be estimated  
297 through  $\omega_{12}=\omega_{13}(\mu_{12}/\mu_{13})$ , with  $\omega$  the peak position and  $\mu$  the reduced mass. We obtain  $\omega_{12}=1593$   
298  $\text{cm}^{-1}$ , which points to a nano-graphitic structure in terms of  $\text{sp}^2$  structure and is entirely consistent

299 with a kerogen structure as found in chondritic IOM, coals and terrestrial kerogens (Ferrari &  
300 Robertson, 2001; Quirico et al., 2018). These data support that most of the unsaturated carbon  
301 atoms are present in small-sized polyaromatic units, with negligible olefinic carbon bonds.

302

### 303 3.3 XANES analysis

304 STXM-based XANES data revealed that the H<sup>13</sup>CO-tholins heated at 400 °C is  
305 homogeneous at the submicrometric level. The prominent feature at 285 – 285.5 eV of the  
306 XANES spectra indicate that more than half of the carbon atoms (56%) are involved in  
307 aromatic/olefinic groups (C=C groups) (Fig. 3a, Table 2). A shoulder is observed at 287.7 eV,  
308 which can be assigned to the aliphatic carbons (-CH<sub>x</sub>) (~ 30% of the C atoms). Additional broad  
309 features at 286.2 eV and 288.5 eV are assigned to ketone/phenols (C=O and C<sub>Ar</sub>-OH, ~ 14 % of  
310 the C atoms) and carboxylic/ester (R-COOH,R'; ~ 4 % of the C atoms). These estimates are  
311 consistent with NMR results. The absorption around 286.5 eV could also be assigned to the  
312 nitrile functional group, but its presence is discarded by the low elemental N/C ratio (0.01)  
313 (Table S-1). Even if our measurements do not report such heterogeneity in H<sup>13</sup>CO-tholins, we  
314 cannot exclude the presence of such tiny inclusions.

315 To validate their XANES quantification method, Le Guillou et al. (2018) applied both  
316 XANES and DP/MAS NMR techniques to fresh lignin and lignin heated at 300 °C and 350 °C  
317 given by Le Brech et al. (2016). This work suggested that XANES and NMR quantification  
318 methods provide consistent results for aromatic groups, ketone/phenol groups and  
319 carboxylic/ester groups. However, the aliphatic groups of lignin samples appeared  
320 overestimated by about 15%. In the case of the heated H<sup>13</sup>CO-tholins, there is a reasonable  
321 agreement between XANES and NMR for the aliphatic and the carboxylic groups (Fig. 3b). In  
322 contrast, the aromatic groups' abundances determined from XANES are lower than that obtained  
323 from DP/MAS NMR, as 55 ± 4 % and 64 ± 1 %, respectively.

324

### 325 3.4 Infrared spectroscopy

326 The infrared spectra of the initial and 400 °C H<sup>12</sup>CO and H<sup>13</sup>CO-tholins are presented in  
327 Fig. 4. The heterogeneity of the 400 °C H<sup>13</sup>CO-tholins was checked in both the KBr pellet and  
328 Micro-FTIR spectra (Supporting Information, Fig. S-4). The broadband at 3600 – 3100 cm<sup>-1</sup>  
329 corresponds to the stretching mode of – OH groups in alcohols, carboxylic groups or phenol  
330 (C<sub>aro</sub>-OH). The peak at 3100 – 3000 cm<sup>-1</sup> is assigned to aromatic C-H. The region of 2800–3000  
331 cm<sup>-1</sup> displays four main peaks, including the antisymmetric and symmetric stretching modes of  
332 methyl (CH<sub>3</sub>) and methylene (CH<sub>2</sub>) groups and possibly of the CH group, as a component  
333 centered at ~ 2900 cm<sup>-1</sup>. However, this spectral range is highly congested and more components  
334 are present (see below). The spectral region 1800 – 1500 cm<sup>-1</sup> displays bands at 1656 and 1540  
335 cm<sup>-1</sup>, and 1708 cm<sup>-1</sup> and 1608 cm<sup>-1</sup> for the H<sup>13</sup>CO and H<sup>12</sup>CO-tholins, respectively. These  
336 features (e.g., 1656 and 1540 cm<sup>-1</sup>) are assigned to the stretching modes of the carbonyl (C=O)  
337 and aromatic (C=C), respectively (Painter et al., 1981). In the region 1500 – 900 cm<sup>-1</sup>, bending,  
338 deformation and out-of-plane modes of CH<sub>2</sub>, CH<sub>3</sub> and C-O bonds are present at 1450, 1375 and  
339 1300 cm<sup>-1</sup>. We will follow here Painter et al. (1981) which state that the broad feature peaking

340 at  $\sim 1200\text{ cm}^{-1}$  does not provide any significant compositional information, in particular on the  
341 abundance of ether groups (-C-O-C-).

342

### 343 **3.4.1 Spectral decomposition of the $\text{CH}_x$ stretching region**

344 The region between  $3000 - 2750\text{ cm}^{-1}$  results from the contribution of symmetric and  
345 antisymmetric stretching modes of  $\text{CH}_2$ ,  $\text{CH}_3$ , possibly of -CH, and contains information on the  
346 abundance of alkyl groups and the ratio of  $\text{CH}_2$  and  $\text{CH}_3$  groups. However, the spectral  
347 congestion makes it difficult to identify the exact number of components, spectral profiles and  
348 physical interpretation. Pure alkanes at a low temperature usually display several components  
349 ascribed to the  $d^-$  antisymmetric stretch of chain  $\text{CH}_2$ , the symmetric stretch of terminal  $\text{CH}_2$   
350 ( $d^+\omega$ ) and two components arising from a Fermi resonance between the symmetric stretch of  
351 chain  $\text{CH}_2$  and the overtone of the H-C-H bending mode, reported as  $d^+$  and  $d^+_{\text{FR}}$  (the first being  
352 dominating by the fundamental mode) (Fig. 5) (MacPhail et al., 1984); other components are  
353 the  $r_a$  and  $r_b$  in-plane and out-of-plane asymmetric stretch of  $-\text{CH}_3$ , and the symmetric stretch  
354 in Fermi resonance with the first overtone of the  $\text{CH}_3$  bending mode, named  $r^+$  and  $r^+_{\text{FR}}$ . Pure  
355 alkanes measured at low-temperature display peaks and shoulders in agreement with this  
356 assignment. However, the different components cannot be guessed for a similar compound in  
357 the liquid state at room temperature (Fig. 5). Even more, critical issues arise in the case of natural  
358 complex mixtures or solids, such as broader components, spectral shapes changes and peak  
359 position shifts. In the case of coals and chondritic IOM, the presence of oxygen as ether or  
360 carbonyl groups likely modifies the spectral distribution, making uncertain the spectral  
361 decomposition of the bands (Orthous-Daunay et al., 2013; Quirico et al., 2018).

362 To explore more quantitatively this congestion issue, we ran optimization simulations of  
363 the fit of the alkyl band of an IOM sample extracted from the Orgueil (CI1) chondrite (Orthous-  
364 Daunay et al., 2013). We used the Artificial Bees algorithm from the ESA Package PYGMO  
365 (<https://esa.github.io/pygmo/>) and tested two fit models: one based on 7 Gaussian and one on  
366 seven pseudo-Voigt components. Two components aimed to fit the small features at  $\sim 3000$  and  
367  $2790\text{ cm}^{-1}$ , due to the aromatic C-H stretch and a combination mode from fundamental modes  
368 in the fingerprint region (Fig. 6). The fit model of the aliphatic bands then consisted of five  
369 components, as usually reported in the literature (Dartois et al., 2007, 2004; Orthous-Daunay et  
370 al., 2013). We ran two groups of simulations: one keeping constant the Full Width at Half  
371 Maximum FWHM of  $35\text{ cm}^{-1}$  for all the components, and another one where FWHM is let free  
372 within the range  $20\text{-}40\text{ cm}^{-1}$ . Each group comprised 50 simulations, and the fit quality was  
373 estimated with a  $\chi^2$  function. A 20000 bees' population over 50 generations were used. For each  
374 of the 200 numerical simulations, the resulting fit was excellent in terms of  $\chi^2$  value (Fig. 6a).  
375 The results are summarized in Fig. 6b-c. Two parameters are of particular interest: (1) the ratio  
376 of the integrated absorbance of the antisymmetric stretching components of  $\text{CH}_2$  and  $\text{CH}_3$   
377 ( $a_{\text{CH}_2}/a_{\text{CH}_3}$ ) and (2) the sum of these values. The first estimates the ratio of the number of  $\text{CH}_2$   
378 over  $\text{CH}_3$  groups, while the second is a measure of the alkyl content ( $a_{\text{CH}_2}+a_{\text{CH}_3}$ ). Note, however,  
379 that for this latter the abundance of the CH group is not measured regarding the assignment of  
380 the central component peaking at  $\sim 2900\text{ cm}^{-1}$  to the overtone of a bending mode in Fermi  
381 resonance with a fundamental C-H stretch (Dartois et al., 2007; 2004). In the case of FWHM

382 forced to  $35\text{ cm}^{-1}$ , we obtain  $1.36 \pm 0.07$  and  $1.52 \pm 0.28$  for  $a_{\text{CH}_2}/a_{\text{CH}_3}$  and  $57.0 \pm 1.8$  and  $59.0$   
383  $\pm 2.7$  for  $a_{\text{CH}_2+a_{\text{CH}_3}}$ , for the Gaussian and pseudo-Voigt models, respectively (Fig. 6b). If FWHM  
384 is let free within  $20\text{--}40\text{ cm}^{-1}$ , we obtain  $1.37 \pm 0.79$  and  $1.66 \pm 0.90$  for  $a_{\text{CH}_2}/a_{\text{CH}_3}$  and  $50.4 \pm 6.3$   
385 and  $52.8 \pm 7$  for  $a_{\text{CH}_2+a_{\text{CH}_3}}$ , for the Gaussian and pseudo-Voigt models, respectively. The  
386 dispersion is much larger, and the absolute values of the sum are also significantly different  
387 from those obtained by forcing FWHM. To make things clear on the impact in terms of chemical  
388 information, we can estimate the  $n_{\text{CH}_2}/n_{\text{CH}_3}$  of Orgueil IOM, using the  $A_{\text{CH}_2}$  and  $A_{\text{CH}_3}$  values  
389 from [Dartois et al. \(2007\)](#). We find that  $n_{\text{CH}_2}/n_{\text{CH}_3}$  ranges between 0.5 and 4.5, while NMR  
390 measurements provide a value ranging 1.9-2.6 ([Gardinier et al. 2000](#); assuming that ether  
391 bridges are  $\text{CH}_3\text{-O-CH}_2\text{:CH}_2\text{-O-CH}_2=1:1$  or  $\text{CH}_2\text{-O-CH}_2$ , respectively). This result is basically  
392 of little value. In practice, things may not be as worse, provided that the fit model is similar to  
393 that used to derive the absorption cross-sections. Finally, due to the sensitivity of the fit models  
394 on derived parameters, and to propose an easy procedure allowing cross-comparison of data  
395 collected by independent groups, we prefer to use the integrated absorbance of the whole alkyl  
396 band, in place of  $a_{\text{CH}_2+a_{\text{CH}_3}}$ . The calibration presented below will lose accuracy when the  
397  $n_{\text{CH}_2}/n_{\text{CH}_3}$  ratio departs from that of the alkane's standards, but as discussed below, the impact  
398 is very limited.

399

#### 400 **3.4.2 Determining the $CH_x/A_{C=C}$ and $A_{CH_x}/A_{C=O}$ ratios**

401 The ratio of the absorption cross-section of two chemical groups  $X_1$  and  $X_2$  is calculated  
402 from the Equation 1. The term  $n_{X_2}/n_{X_1}$  is provided by the NMR analysis, and the integrated  
403 absorbances of the alkyl, carbonyl and aromatic groups are measured in the spectra of the  $400$   
404  $^\circ\text{C H}^{13}\text{CO}$ -tholins. The uncertainty values were estimated by considering fit uncertainties, the  
405 baseline correction, sample heterogeneity (Fig. S-4) and the uncertainty on NMR values of the  
406 relative abundance of the three chemical groups. We finally obtain the ratio values of integrated  
407 cross-section:  $A_{\text{CH}_x}/A_{\text{Aro}} = 1.578 \pm 0.183$  and  $A_{\text{CH}_x}/A_{\text{C=O}} = 0.116 \pm 0.006$  (Table 3).

408

## 409 **4 DISCUSSION**

### 410 **4.1 Achievements and limitations**

411 The composition of the  $\text{H}^{13}\text{CO}$ -tholin determined by NMR is very reliable for the  
412 aliphatic and aromatic contents, due to intense DP/MAS features at 38 and 130 ppm that are  
413 unambiguously assigned to alkyl and conjugated carbons. Unlike CP/MAS data, these spectra  
414 do not suffer from selective signal enhancement and are suitable for quantification. The  
415 assignment is also robust for oxygen-bearing groups, but the features at 169 and 200 ppm are  
416 weak and suffer from a higher uncertainty in their relative abundances. The new calibration of  
417  $\text{H}^{13}\text{CO}$ -tholin allows determining the absorption cross-section ratios of the  $\text{CH}_x$ ,  $\text{C=O}$  and Aro  
418 groups.

419 Nevertheless, the calibration derived from the  $\text{H}^{13}\text{CO}$ -tholins suffers from limitations  
420 and biases that are intrinsic to infrared spectroscopy. First, as we estimate the alkyl abundance  
421 through the integration of all the alkyl bands ( $\text{CH}_x$  stretching in the range of  $3000\text{--}2800\text{ cm}^{-1}$ ),

422 the validity of the H<sup>13</sup>CO-tholins as a calibration standard relies on a similar distribution of the  
 423 alkyl groups in the kerogen and coal samples that are analyzed. Hence, the absorption cross-  
 424 section of CH<sub>2</sub> and CH<sub>3</sub> are not identical, e.g. for the anti-symmetric stretching modes A<sub>CH<sub>3</sub></sub> =  
 425 (11.8 ± 0.76) × 10<sup>-18</sup> and A<sub>CH<sub>2</sub></sub> = (7.7 ± 1.5) × 10<sup>-18</sup> cm/group (Dartois et al., 2007; Sandford et  
 426 al., 1991). Considering a variation range of 1.5 – 3.0 for n<sub>CH<sub>2</sub></sub>/n<sub>CH<sub>3</sub></sub> that is observed in chondritic  
 427 IOM (Orthous-Daunay et al., 2013), we can calculate the variation of the sum of the integrated  
 428 absorbance of the anti-symmetric stretching modes of CH<sub>2</sub> and CH<sub>3</sub>, a=a<sub>CH<sub>2</sub></sub>+a<sub>CH<sub>3</sub></sub>, as:

$$429 \quad \frac{a_{1.5}}{a_3} = \frac{1 + \frac{1}{1.5} \frac{A_{CH_3}}{A_{CH_2}}}{1 + \frac{1}{3} \frac{A_{CH_3}}{A_{CH_2}}} = 1.18$$

430 This means that n<sub>CH<sub>2</sub></sub>/n<sub>CH<sub>3</sub></sub> estimates on IOM from thermally processed chondrites  
 431 obtained with our calibration are underestimated by 20% at maximum. Second, a critical issue  
 432 with the 1600 cm<sup>-1</sup> band is the sensitivity of its absorption cross-section to the oxygen molecular  
 433 environment, in particular the phenolic (-OH) and possibly the ether (C-O-C) groups (Painter et  
 434 al., 1981). This means that the kerogen to be analyzed should be as similar as possible to the  
 435 calibration standard in terms of oxygen speciation around aromatic carbons. In this respect, the  
 436 400°C H<sup>13</sup>CO-tholins do not fit mature samples that have large polyaromatic units, and which  
 437 are known to have a lower aromatic cross-section. This issue is probably more critical for coals,  
 438 due to the large O/C variations upon maturity.

439

## 440 4.2 IOM composition of primitive chondrites

441 A way to test our calibration is to apply our method to samples whose composition has  
 442 been previously determined by NMR. We report here the composition (n<sub>CH<sub>x</sub></sub>/n<sub>C=O</sub> and n<sub>CH<sub>x</sub></sub>/n<sub>C=C</sub>)  
 443 of the IOMs from the Orgueil, Murchison, EET 92042 and Tagish Lake chondrites, derived  
 444 from the spectra published by Orthous-Daunay et al. (2013) (Fig. S-5). These compositions are  
 445 compared with the NMR estimates of Cody & Alexander, (2005) and Gardinier et al. (2000)  
 446 (Fig. 7). It is first necessary to discuss those NMR data, which display some inconsistencies  
 447 between them.

448 Two types of NMR measurements are presented in these publications: Single-pulse (SP)  
 449 MAS and Cross Polarized (CP) MAS. Single-pulse (SP) MAS NMR is usually presented as  
 450 suitable for quantification (Baccile et al., 2014), but in the case of natural samples (<sup>13</sup>C/<sup>12</sup>C ~  
 451 1%), the signal is low, long acquisition time is required, and eventually the spectra still contain  
 452 significant noise. In the case of broad features, band distortion can happen due to dephasing  
 453 processes before spectrum acquisition, leading to significant errors in the quantification of band  
 454 integral. In addition, the baseline may depart from an ideal flat line, thereby perturbing the  
 455 profile of broad features. Altogether, these artifacts may lead to significant errors. These issues  
 456 can be fixed, at least partly, by using spin echo DP/MAS experimental configuration, as done  
 457 here on the standard samples. By applying two subsequent pulses (P<sub>90-τ</sub>-P<sub>180-τ</sub>, with τ being a  
 458 multiple of the MAS frequency), a rephasing of the signal is obtained, and broad features are  
 459 not anymore distorted (Hahn, 1950; Vieira et al., 2016; Lucherelli et al., 2019). Cross  
 460 polarization Magic Angle Spinning (CP/MAS) analysis is based on the magnetization transfer



461 from hydrogen to carbon atoms. It provides spectra with an improved signal-to-noise ratio, but  
462 the quantification is not straightforward (Baccile et al., 2014; dela Rosa et al., 1992). Thus, the  
463 intensity of the signal arising from a carbon atom depends on its hydrogen environment, and the  
464 signal depends on the contact time. It requires an estimate of the fraction of undetected carbon  
465 atoms through calculations constrained by the elemental composition of the sample. The  
466 CP/MAS estimation is then much less direct than DP/MAS, and may suffer from error  
467 propagation on speciation quantitation and elemental composition, as well as a non-  
468 comprehensive identification of all involved carbon speciation.

469 Cody and Alexander, (2005) and Gardinier et al. (2000) both report CP/MAS NMR data  
470 that provide quantitative estimates of the composition of IOM from Orgueil and Murchison. In  
471 the case of the Murchison chondrite, the ratio of aliphatic to aromatic carbon atoms  $F_{\text{ali}}/F_{\text{aro}}$   
472 estimated by Gardinier et al. (2000) is slightly higher than that of Cody and Alexander (2005)  
473 but it lies within the error bars (Fig. 8). In the case of Orgueil, the  $F_{\text{ali}}/F_{\text{aro}}$  ratio from is lower  
474 than that of Cody and Alexander (2005), but the error bars also overlap (Fig. 8). In contrast,  
475 there is some significant disagreement regarding carbon atoms in oxygenated functions, in  
476 particular for Orgueil (Fig. 8). These differences may be due to different experimental  
477 procedures. Gardinier et al. (2000) investigated the spectral dependence upon contact time,  
478 applied a correction accounting for undetected carbon atoms based on the bulk H/C ratio,  
479 determined from the combustion experiment. The abundance of undetected aromatic carbon was  
480 found significant, ranging from 20 – 29%. We note that the H/C value (0.7 in their study is  
481 higher than that measured by Alexander et al. (2007, 2013) ( $0.588 \pm 0.016$  and  $0.618$ ,  
482 respectively), which may have an impact on the derived abundances. Cody and Alexander  
483 (2005) did not run systematic measurements as a function of contact time, but this latter was set  
484 to 4.5 ms to fit with SP/MAS spectra. This quantitative interpretation of CP/MAS spectra  
485 without investigations at variable contact time is not usual, and estimates might suffer from  
486 biases (S. Derenne, personal communication; see also). In fact, in the details, the estimates  
487 derived by Cody and Alexander (2005) from SP/MAS and CP/MAS show significant  
488 differences. Due to these complex issues with CP/MAS, we have restricted to the SP/MAS data  
489 from Cody and Alexander (2005) for comparison with our FTIR data.

490 We observe that  $n_{\text{CH}_x}/n_{\text{C}=\text{C}}$  estimated from infrared data following equation 2 fits well  
491 NMR data in the case of Orgueil, Murchison and Tagish Lake.  $n_{\text{CH}_x}/n_{\text{C}=\text{C}}$  in EET 92042 is  
492 underestimated by a factor of  $\sim 1.3$ , though the error bars overlap (Fig. 8). The first explanation  
493 of this moderate discrepancy is the chemical control of the  $A_{\text{Aro}}$  absorption cross-section by  
494 oxygenated groups (Painter et al., 1981). The O/C ratio in EET 92042 ( $\text{O/C}=16.6 \pm 0.2$ ) is lower  
495 than those of Murchison ( $18.3 \pm 0.3$ ) and Orgueil ( $18.1 \pm 0.2$ ). These values are quite similar.  
496 According to NMR data, the main oxygen speciation in chondritic IOM is carbonyl (ketone,  
497 ester) and ether bridges. The estimations from SP/MAS NMR spectra are challenging due to the  
498 faintness of the features and the poor signal-to-noise ratio, and they do not point to significant  
499 variations between EET 92042, and Orgueil and Murchison. Altogether, there is no evidence of  
500 a chemical control of the  $A_{\text{Aro}}$  absorption cross-section. A second tentative explanation is that  
501 the SP/MAS data are not quantitative enough, due to the biases abovementioned. The only way  
502 to firmly solve this issue would be to collect DP/MAS NMR spectra (MAS synchronized spin  
503 echoes) of IOMs from those three chondrites.

504 Regarding carbonylated species, we observe that  $n_{\text{CH}_x}/n_{\text{C=O}}$  is systematically higher than  
505 NMR estimates, though the hierarchy Orgueil > Murchison > Tagish Lake is similar to NMR  
506 estimates (Fig. 8). The number of bonds is equal to the number of carbon atoms in the case of  
507  $\text{CH}_x$  groups and carbonyl; therefore, the difference may be due here to different absorption  
508 cross-sections of the  $\text{C=O}$  group between these IOMs and the HCO-tholin standard. As for the  
509  $\text{C=C}$  bonds, the local molecular environment plays a role in the control of the absorption cross-  
510 section of the  $\text{C=O}$  stretching mode. Such variations have been observed for small molecules in  
511 the solid-state, for instance, amorphous ices of formic acid and acetaldehyde, whose carbonyl  
512 absorption cross-sections differ by a factor of 2 (Bouilloud et al., 2015; Hudson & Ferrante,  
513 2020). Nevertheless, as mentioned in the previous section, the re-analysis of these IOM samples  
514 by DP/MAS NMR (MAS synchronized spin echoes) is required to get a firm conclusion.

515 *To sum up, the  $\text{H}^{13}\text{CO}$ -tholins standard provides a suitable calibration for estimating*  
516  *$n_{\text{CH}_x}/n_{\text{C=C}}$  in IOMs derived from FTIR measurement compared with SP/MAS NMR. In contrast,*  
517  *$n_{\text{CH}_x}/n_{\text{C=O}}$  is overestimated in all classes of chondrites compared with the SP/MAS NMR from*  
518 *Cody and Alexander (2005).*

519

### 520 4.3 Composition of ligno-cellulosic coals

521 We report now the composition of a series of coals of varying ranks from lignite to semi-  
522 anthracite (Table S-2, Fig. S-5). The majority of coals plot within the Type III kerogen band on  
523 the “Van Krevelen” diagram (Fig. 9a). Due to spectral congestion, the fit in the region 1800-  
524 1500  $\text{cm}^{-1}$  region is more complicated than for IOMs. We used four spectral components due to  
525 ketones/quinone, aromatic branched with phenolic -OH (or -O-), and aromatic with a carboxylic  
526 group (Chen et al., 2012; Painter et al., 1981; Petersen et al., 2008) (Fig. S-6). The abundance  
527 of aromatic groups is then estimated through the two last components, and the  $\text{C=O}$  by the first  
528 two components. Our results of the evolution trends are consistent with previous estimations  
529 based on integrated abundance derived from IR spectra (Chen et al., 2012). No NMR data are  
530 available for our coals, but NMR data are available for Chinese series of coals (Wei & Tang,  
531 2018) and for the so-called Argonne Premium coals, a subset of the Penn State Coals Data bank  
532 (de la Rosa et al., 1992). Figure 9b-c reports the FTIR and NMR CP/MAS estimates of Wei and  
533 Tang (2018) plotted against the vitrinite reflectance. We observe two opposite trends: while the  
534 NMR estimates point to a progressive decrease of  $n_{\text{CH}_x}/n_{\text{C=C}}$ , the FTIR estimates show an  
535 increase up to the vitrinite reflectance of 1.73% and then a decline for the semi-anthracite  
536 samples. In contrast,  $n_{\text{CO}}/n_{\text{CH}_x}$  is consistent with NMR. This situation is opposite to what he had  
537 with IOMs (Fig. 8).

538 The NMR estimates of Wei and Tang (2018) are, at least qualitatively, in agreement with  
539 the progressive aromatization of the coals during diagenesis and catagenesis (Chen et al., 2012).  
540 This supports the view that the ratio of the cross-section  $A_{\text{CH}_x}/A_{\text{Aro}}$  varies with the vitrinite  
541 reflectance, as previously reported (Painter et al., 1981; Dela Rosa et al., 1992). It also shows  
542 that any semi-quantitative analysis based on IR measurements leads to inaccurate estimates of  
543 the chemical evolution. NMR data collected by Wei & Tang (2018) were obtained in CP/MAS  
544 with a 3 ms contact time. The publication does not report measurements over variable time



545 contact, nor the acquisition of single pulse spectra. In contrast, [Dela Rosa et al. \(1992\)](#) report  
546 varying contact time CP/MAS measurements on Argonne Premium coals and one single pulse  
547 (Bloch decay) analysis of Pittsburgh n°8 that fits the corresponding CP/MAS estimate. Muntean  
548 et al. (1990) also report consistent CP/MAS and single pulse estimates in the case of Pocahontas  
549 N°3, provided that the delay time remains < 75 μs. However, the agreement between both types  
550 of measurements may be worse in the case of O-rich low-rank coals ([Dela Rosa et al., 1992](#)).  
551 We have recalculated the aromaticity of the coals from [Wei and Tang \(2018\)](#) following the  
552 definition of [Dela Rosa et al. \(1992\)](#), and found that the difference between estimates of the two  
553 sets of data does not exceed 5%.

554 Finally, using those data and keeping in mind that uncertainties might be worse for the  
555 more immature coals, we have calculated the cross-section ratio  $A_{CHx}/A_{C=C}$  upon the vitrinite  
556 reflectance (Fig. 9d). This curve is best fitted with a 6<sup>th</sup> order polynomial (Equation 3):

$$\frac{A_{CHx}}{A_{C=C}} = \sum_{i=0}^6 K_i (x - x_0)^i \quad (\text{Equation 3})$$

557  
558 with  $K_0=0.64599$ ;  $K_1=0.23482$ ;  $K_2=6.3733$ ;  $K_3=-16.939$ ;  $K_4=23.446$ ;  $K_5=-14.846$ ;  
559  $K_6=3.7045$  and  $X_0=0.3$ . This formula applies in the range  $R_o = 0.3 - 2\%$ , provided that a similar  
560 fit model is used to decompose the coal infrared spectra. We observe that  $A_{CHx}/A_{C=C}$  increases  
561 and  $A_{C=C}$  decreases, upon increasing the vitrinite reflectance  $R_o$ , consistently with control of the  
562 absorption cross-section by oxygenated groups. The increase is quite high, around one order of  
563 magnitude, which means that semi-quantitative sets of infrared data must be imperatively  
564 corrected. However, further investigations combining DP/MAS NMR and FTIR measurements  
565 on the same series of coals should provide more accurate data.

566

## 567 **5 CONCLUSION**

568 We have tested a protocol for quantifying the abundance of the CH<sub>x</sub>, C=O and C=C  
569 groups in chondritic IOMs and coals, based on a <sup>13</sup>C model material characterized by DP/MAS  
570 NMR and FTIR. The ratios of the cross-sections of these groups have been determined, and the  
571 method has been applied to IOM samples extracted from CI, CM and CR chondrites, and to a  
572 series of coals. The estimate of  $n_{CHx}/n_{C=C}$  in IOM from CI and CM chondrites is fair, but  
573  $n_{CHx}/n_{C=O}$  is overestimated. The estimate of  $n_{CHx}/n_{C=C}$  in the CR chondrite is lower than the  
574 estimation by SP/MAS published earlier. These discrepancies are due to the control of the  
575 absorption of the C=O and C=C bonds by the local oxygenated chemical structure. In this regard,  
576 the oxygen speciation in CR chondrites may be significantly different than that of CI and CM  
577 chondrites. However, new NMR analyses of IOM from CR chondrites are necessary to test the  
578 reproducibility of existing data. Regarding coals, the use of published NMR analyses shows that  
579 the ratio  $A_{CHx}/A_{C=C}$  varies with the vitrinite reflectance over an order of magnitude. Here as  
580 well, the local oxygen speciation plays a critical control on  $A_{Aro}$ , which decreases with  
581 increasing the vitrinite reflectance. We propose an analytical law for  $A_{CHx}/A_{Aro}$ , which should  
582 be helpful to determine the aromaticity of any coal sample, provided its vitrinite reflectance is  
583 known.



## 585 ACKNOWLEDGEMENTS

586 The authors are grateful to the financial support of the H2020 European Research Council (ERC)  
587 (SOLARYS ERC-CoG2017\_771691). We acknowledge Dr. Julien Alleon for the assistance  
588 during STXM-XANES measurements at the HERMES beamline at the SOLEIL Synchrotron.  
589 We also acknowledge Anne-Marie Blanchenet who prepared the microtomed section for STXM  
590 analysis at the microscopy platform of the Lille University. We are thankful to the Raman  
591 facility in ENS-Lyon, supported by the Institut National des Sciences de l'Univers (INSU). We  
592 also thank Paul Craddock and an anonymous reviewer for helpful comments that improved the  
593 manuscript.

594

## 595 Supplementary data

596 Supplementary data related to this article can be found at the Supporting Information part.

597

## 598 Reference

599 Abarghani, A., Gentzis, T., Shokouhimehr, M., Liu, B., & Ostadhassan, M. (2020). Chemical  
600 heterogeneity of organic matter at nanoscale by AFM-based IR spectroscopy. *Fuel*,  
601 261(116454). <https://doi.org/10.1016/j.fuel.2019.116454>

602 Alcouffe, G., Cavarroc, M., Cernogora, G., Ouni, F., Jolly, A., Boufendi, L., & Szopa, C.  
603 (2010). Capacitively coupled plasma used to simulate Titan's atmospheric chemistry.  
604 *Plasma Sources Science and Technology*, 19(1). [https://doi.org/10.1088/0963-](https://doi.org/10.1088/0963-0252/19/1/015008)  
605 0252/19/1/015008

606 Baccile, N., Falco, C., & Titirici, M. M. (2014). Characterization of biomass and its derived  
607 char using <sup>13</sup>C-solid state nuclear magnetic resonance. *Green Chemistry*, 16(12), 4839–  
608 4869. <https://doi.org/10.1039/c3gc42570c>

609 Battandier, M., Bonal, L., Quirico, E., Beck, P., Engrand, C., Duprat, J., & Dartois, E. (2018).  
610 Characterization of the organic matter and hydration state of Antarctic micrometeorites:  
611 A reservoir distinct from carbonaceous chondrites. *Icarus*, 306, 74–93.  
612 <https://doi.org/10.1016/j.icarus.2018.02.002>

613 Beck, P., Quirico, E., Montes-hernandez, G., L., B., J., B., Orthous-Daunay, F.-R., ... O., B.  
614 (2010). Hydrous mineralogy of CM and CI chondrites from infrared spectroscopy and  
615 their relationship with low Albedo asteroid. In *41st Lunar and Planetary Science*.

616 Belkhou, R., Stanescu, S., Swaraj, S., Besson, A., Ledoux, M., Hajlaoui, M., & Dalle, D.  
617 (2015). HERMES: A soft X-ray beamline dedicated to X-ray microscopy. *Journal of*  
618 *Synchrotron Radiation*, 22, 968–979. <https://doi.org/10.1107/S1600577515007778>

619 Bouilloud, M., Fray, N., Bénilan, Y., Cottin, H., Gazeau, M. C., & Jolly, A. (2015).  
620 Bibliographic review and new measurements of the infrared band strengths of pure  
621 molecules at 25 K: H<sub>2</sub>O, CO<sub>2</sub>, CO, CH<sub>4</sub>, NH<sub>3</sub>, CH<sub>3</sub>OH, HCOOH and H<sub>2</sub>CO. *Monthly*  
622 *Notices of the Royal Astronomical Society*, 451(2), 2145–2160.  
623 <https://doi.org/10.1093/mnras/stv1021>

624 Chen, Y., Mastalerz, M., & Schimmelmann, A. (2012). Characterization of chemical  
625 functional groups in macerals across different coal ranks via micro-FTIR spectroscopy.  
626 *International Journal of Coal Geology*, 104, 22–33.  
627 <https://doi.org/10.1016/j.coal.2012.09.001>

- 628 Cody, G. D., & Alexander, C. M. O. D. (2005). NMR studies of chemical structural variation  
629 of insoluble organic matter from different carbonaceous chondrite groups. *Geochimica et*  
630 *Cosmochimica Acta*, 69(4), 1085–1097. <https://doi.org/10.1016/j.gca.2004.08.031>
- 631 Craddock, P. R., Doan, T. V. Le, Bake, K., Polyakov, M., Charsky, A. M., & Pomerantz, A.  
632 E. (2015). Evolution of Kerogen and Bitumen during Thermal Maturation via Semi-  
633 Open Pyrolysis Investigated by Infrared Spectroscopy. *Energy and Fuels*, 29, 2197–  
634 2210. <https://doi.org/10.1021/ef5027532>
- 635 Dartois, E., Geballe, T. R., Pino, T., Cao, A.-T., Jones, A., Deboffle, D., ... D'Hendecourt, L.  
636 (2007). IRAS 08572+3915: Constraining the aromatic versus aliphatic content of  
637 interstellar HACs. *Astronomy & Astrophysics*, 463, 635–640.  
638 <https://doi.org/10.1051/0004-6361>
- 639 Dartois, E., Munoz Caro, G. M., Deboffle, D., & D'Hendecourt, L. (2004). Diffuse  
640 interstellar medium organic polymers Photoproduction of the 3.4, 6.85 and 7.25  $\mu\text{m}$   
641 features. *Astronomy & Astrophysics*, 423, 33–36.
- 642 Dela Rosa, L., Pruski, M., Lang, D., Gerstein, B., & Solomon, P. (1992). Characterization of  
643 the Argonne Premium Coals by using  $^1\text{H}$  and  $^{13}\text{C}$  NMR and FT-IR spectroscopies.  
644 *Energy and Fuels*, 6(4), 460–468. <https://doi.org/10.1021/ef00034a016>
- 645 Ferrari, A. C., & Robertson, J. (2001). Resonant Raman spectroscopy of disordered,  
646 amorphous, and diamondlike carbon. *Physical Review B - Condensed Matter and*  
647 *Materials Physics*, 64(7), 1–13. <https://doi.org/10.1103/PhysRevB.64.075414>
- 648 Fung, B. M., Khitrin, A. K., & Ermolaev, K. (2000). An Improved Broadband Decoupling  
649 Sequence for Liquid Crystals and Solids. *Journal of Magnetic Resonance*, 142(1), 97–  
650 101. <https://doi.org/10.1006/jmre.1999.1896>
- 651 Gardinier, a, Derenne, S., Robert, F., Behar, F., Largeau, C., & Maquet, J. (2000). Solid state  
652 CP/MAS  $^{13}\text{C}$  NMR of the insoluble organic matter of the Orgueil and Murchison  
653 meteorites: quantitative study. *Earth and Planetary Science Letters*, 184, 9–21.  
654 [https://doi.org/http://dx.doi.org/10.1016/S0012-821X\(00\)00300-9](https://doi.org/http://dx.doi.org/10.1016/S0012-821X(00)00300-9)
- 655 Guillou, C. Le, Bernard, S., Pena, F. De, & Brech, Y. Le. (2018). XANES-Based  
656 Quantification of Carbon Functional Group Concentrations. *Analytical Chemistry*, 90,  
657 8379–8386. research-article. <https://doi.org/10.1021/acs.analchem.8b00689>
- 658 Hahn, E. L. (1950). Spin Echoes. *Phys Rev Lett*.
- 659 Hudson, R. L., & Ferrante, R. F. (2020). Quantifying acetaldehyde in astronomical ices and  
660 laboratory analogues: IR spectra, intensities,  $^{13}\text{C}$  shifts, and radiation chemistry.  
661 *Monthly Notices of the Royal Astronomical Society*, 492, 283–293.  
662 <https://doi.org/10.1093/mnras/stz3323>
- 663 Jakob, D. S., Wang, L., Wang, H., & Xu, X. G. (2019). Spectro-Mechanical Characterizations  
664 of Kerogen Heterogeneity and Mechanical Properties of Source Rocks at 6 nm Spatial  
665 Resolution. *Analytical Chemistry*, 91(14), 8883–8890.  
666 <https://doi.org/10.1021/acs.analchem.9b00264>
- 667 Jubb, A. M., Birdwell, J. E., Hackley, P. C., Hatcherian, J. J., & Qu, J. (2020). Nanoscale  
668 Molecular Composition of Solid Bitumen from the Eagle Ford Group across a Natural  
669 Thermal Maturity Gradient. *Energy and Fuels*, 34(7), 8167–8177.  
670 <https://doi.org/10.1021/acs.energyfuels.0c00963>
- 671 Kebukawa, Y., Alexander, C. M. O. D., & Cody, G. D. (2011). Compositional diversity in  
672 insoluble organic matter in type 1, 2 and 3 chondrites as detected by infrared  
673 spectroscopy. *Geochimica et Cosmochimica Acta*, 75(12), 3530–3541.

- 674 <https://doi.org/10.1016/j.gca.2011.03.037>
- 675 Kebukawa, Y., Kobayashi, H., Urayama, N., Baden, N., & Kondo, M. (2018). Nanoscale  
676 infrared imaging analysis of carbonaceous chondrites to understand organic-mineral  
677 interactions during aqueous alteration. *PNAS*, 1–6.  
678 <https://doi.org/10.1073/pnas.1816265116>
- 679 Kingsley, P. B. (1999). Methods of measuring spin-lattice (T1) relaxation times: An  
680 annotated bibliography. *Concepts in Magnetic Resonance*, 11(4), 243–276.  
681 [https://doi.org/10.1002/\(SICI\)1099-0534\(1999\)11:4<243::AID-CMR5>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1099-0534(1999)11:4<243::AID-CMR5>3.0.CO;2-C)
- 682 Le Brech, Y., Raya, J., Delmotte, L., Brosse, N., Gadiou, R., & Dufour, A. (2016).  
683 Characterization of biomass char formation investigated by advanced solid state NMR.  
684 *Carbon*, 108, 165–177. <https://doi.org/10.1016/j.jrp.2009.07.007>
- 685 Le Guillou, C., Bernard, S., De La Pena, F., & Le Brech, Y. (2018). XANES-Based  
686 Quantification of Carbon Functional Group Concentrations. *Analytical Chemistry*,  
687 90(14), 8379–8386. <https://doi.org/10.1021/acs.analchem.8b00689>
- 688 Li, W., Zhu, Y. M., Wang, G., & Jiang, B. (2016). Characterization of coalification jumps  
689 during high rank coal chemical structure evolution. *Fuel*, 185, 298–304.  
690 <https://doi.org/10.1016/j.fuel.2016.07.121>
- 691 Lis, G. P., Mastalerz, M., Schimmelmann, A., Lewan, M. D., & Stankiewicz, B. A. (2005).  
692 FTIR absorption indices for thermal maturity in comparison with vitrinite reflectance R0  
693 in type-II kerogens from Devonian black shales. *Organic Geochemistry*, 36(11), 1533–  
694 1552. <https://doi.org/10.1016/j.orggeochem.2005.07.001>
- 695 Lucherelli, M. A., Raya, J., Edelthalhammer, K. F., Hauke, F., Hirsch, A., Abellán, G., &  
696 Bianco, A. (2019). A Straightforward Approach to Multifunctional Graphene. *Chemistry*  
697 *- A European Journal*, 25(57), 13218–13223. <https://doi.org/10.1002/chem.201903165>
- 698 MacPhail, R. A., Strauss, H. L., Snyder, R. G., & Eiliger, C. A. (1984). C-H stretching modes  
699 and the structure of n-alkyl chains. 2. Long, all-trans chains. *Journal of Physical*  
700 *Chemistry*, 88(3), 334–341. <https://doi.org/10.1021/j150647a002>
- 701 Mapelli, C., Castiglioni, C., Zerbi, G., & Mullen, K. (1999). Common force field for graphite  
702 and polycyclic aromatic hydrocarbons. *Physical Review B - Condensed Matter and*  
703 *Materials Physics*, 60(18), 12710–12725. <https://doi.org/10.1103/PhysRevB.60.12710>
- 704 Massiot, D., Fayon, F., Capron, M., King, I., Le Calvé, S., Alonso, B., ... Hoatson, G. (2002).  
705 Modelling one- and two-dimensional solid-state NMR spectra. *Magnetic Resonance in*  
706 *Chemistry*, 40(1), 70–76. <https://doi.org/10.1002/mrc.984>
- 707 Mathurin, J., Dartois, E., Pino, T., Engrand, C., Duprat, J., Deniset-besseau, A., ... Dazzi, A.  
708 (2019). Nanometre-scale infrared chemical imaging of organic matter in ultra-  
709 carbonaceous Antarctic micrometeorites ( UCAMMs ). *Astronomy & Astrophysics*, 160,  
710 1–9.
- 711 Orthous-Daunay, F. R., Quirico, E., Beck, P., Brissaud, O., Dartois, E., Pino, T., & Schmitt,  
712 B. (2013). Mid-infrared study of the molecular structure variability of insoluble organic  
713 matter from primitive chondrites. *Icarus*, 223, 534–543.  
714 <https://doi.org/10.1016/j.icarus.2013.01.003>
- 715 Painter, P. P., Snyder, R. W., Starsinic, M., Coleman, M. M., Kuehn, D. W., & Davis, A.  
716 (1981). Concerning the application of FT-IR to the study of coal: A critical assessment  
717 of band assignments and the application of spectral analysis programs. *Applied*  
718 *Spectroscopy*, 35(5), 475–485. <https://doi.org/10.1366/0003702814732256>
- 719 Petersen, H. I., Rosenberg, P., & Nytoft, H. P. (2008). Oxygen groups in coals and alginite-

- 720 rich kerogen revisited. *Coal Geology*, 74, 93–113.  
721 <https://doi.org/10.1016/j.coal.2007.11.007>
- 722 Pretsch, E., Bühlmann, P., & Affolter, C. (2000). *Structure Determination of Organic*  
723 *Compounds. Structure Determination of Organic Compounds*. Springer.  
724 <https://doi.org/10.1007/978-3-662-04201-4>
- 725 Quirico, E., Bonal, L., Beck, P., Alexander, C. M. O. D., Yabuta, H., Nakamura, T., ...  
726 Schmitt-kopplin, P. (2018). Prevalence and nature of heating processes in CM and C2-  
727 ungrouped chondrites as revealed by insoluble organic matter. *Geochimica et*  
728 *Cosmochimica Acta*, 241, 17–37. <https://doi.org/10.1016/j.gca.2018.08.029>
- 729 Quirico, Eric, Orthous-Daunay, F. R., Beck, P., Bonal, L., Brunetto, R., Dartois, E., ...  
730 Duprat, J. (2014). Origin of insoluble organic matter in type 1 and 2 chondrites: New  
731 clues, new questions. *Geochimica et Cosmochimica Acta*, 136, 80–99.  
732 <https://doi.org/10.1016/j.gca.2014.03.025>
- 733 Russell, L. M., Takahama, S., Liu, S., Hawkins, L. N., Covert, D. S., Quinn, P. K., & Bates,  
734 T. S. (2009). Oxygenated fraction and mass of organic aerosol from direct emission and  
735 atmospheric processing measured on the R/V Ronald Brown during  
736 TEXAQS/GoMACCS 2006. *Journal of Geophysical Research Atmospheres*, 114(7), 1–  
737 15. <https://doi.org/10.1029/2008JD011275>
- 738 Sandford, S. A., Allamandola, L. J., Tielens, A. G. G. M., Sellgren, K., Tapia, M., &  
739 Pendleton, Y. (1991). The interstellar C-H stretching band near 3.4 microns: constraints  
740 on the composition of organic material in the diffuse interstellar medium. *The*  
741 *Astrophysical Journal*, 371, 607–620.
- 742 Sciamma-O'Brien, E., Carrasco, N., Szopa, C., Buch, A., & Cernogora, G. (2010). Titan's  
743 atmosphere: An optimal gas mixture for aerosol production? *Icarus*, 209(2), 704–714.  
744 <https://doi.org/10.1016/j.icarus.2010.04.009>
- 745 Sun, X. (2005). The investigation of chemical structure of coal macerals via transmitted-light  
746 FT-IR microspectroscopy. *Spectrochimica Acta - Part A: Molecular and Biomolecular*  
747 *Spectroscopy*, 62, 557–564. <https://doi.org/10.1016/j.saa.2005.01.020>
- 748 Swaraj, S., Belkhou, R., Stanescu, S., Rioult, M., Besson, A., & Hitchcock, A. P. (2017).  
749 Performance of the HERMES beamline at the carbon K-edge. *Journal of Physics:*  
750 *Conference Series*, 849(1). <https://doi.org/10.1088/1742-6596/849/1/012046>
- 751 Szopa, C., Cernogora, G., Boufendi, L., Correia, J. J., & Coll, P. (2006). PAMPRE : A dusty  
752 plasma experiment for Titan's tholins production and study. *Planetary and Space*  
753 *Science*, 54, 394–404. <https://doi.org/10.1016/j.pss.2005.12.012>
- 754 Takahama, S., Johnson, A., Russell, L. M., Takahama, S., & Johnson, A. (2013).  
755 Quantification of Carboxylic and Carbonyl Functional Groups in Organic Aerosol  
756 Infrared Absorbance Spectra. *Aerosol Science and Technology*, 47(3), 310–325.  
757 <https://doi.org/10.1080/02786826.2012.752065>
- 758 Vieira, M. A., Gonçalves, G. R., Cipriano, D. F., Schettino, M. A., Silva Filho, E. A., Cunha,  
759 A. G., ... Freitas, J. C. C. (2016). Synthesis of graphite oxide from milled graphite  
760 studied by solid-state <sup>13</sup>C nuclear magnetic resonance. *Carbon*, 98, 496–503.  
761 <https://doi.org/10.1016/j.carbon.2015.11.037>
- 762 Wang, J., Morin, C., Li, L., Hitchcock, A. P., Scholl, A., & Doran, A. (2009). Radiation  
763 damage in soft X-ray microscopy. *Journal of Electron Spectroscopy and Related*  
764 *Phenomena*, 170(1–3), 25–36. <https://doi.org/10.1016/j.elspec.2008.01.002>
- 765 Wei, Q., & Tang, Y. (2018). <sup>13</sup>C-NMR study on structure evolution characteristics of high-

- 766 organic-sulfur coals from typical chinese areas. *Minerals*, 8(2).  
767 <https://doi.org/10.3390/min8020049>
- 768 Xueqiu, H., Xianfeng, L., Baisheng, N., & Dazhao, S. (2017). FTIR and Raman spectroscopy  
769 characterization of functional groups in various rank coals. *Fuel*, 206, 555–563.  
770 <https://doi.org/10.1016/j.fuel.2017.05.101>
- 771 Yabuta, H., Naraoka, H., Sakanishi, K., & Kawashima, H. (2005). Solid-state <sup>13</sup>C NMR  
772 characterization of insoluble organic matter from Antarctic CM2 chondrites: Evaluation  
773 of the meteoritic alteration level. *Meteoritics and Planetary Science*, 40(5), 779–787.  
774 <https://doi.org/10.1111/j.1945-5100.2005.tb00979.x>
- 775 Yang, J., Hatcherian, J., Hackley, P. C., & Pomerantz, A. E. (2017). Nanoscale geochemical  
776 and geomechanical characterization of organic matter in shale. *Nature Communications*,  
777 8(2179). <https://doi.org/10.1038/s41467-017-02254-0>
- 778  
779



780 Table 1. Chemical shifts, corrected relative abundances of the different functional carbon  
781 groups in the initial and heated H<sup>13</sup>CO-tholins derived from Figure 1b-c.

782 Table 2. The functional carbon groups concentrations given by the XANES-based  
783 quantification (Le Guillou et al., 2018).

784 Table 3. Integrated cross-section ratio calculation of the 400 °C H<sup>13</sup>CO-tholins based on the  
785 quantitative NMR and the infrared spectral decomposition.

786

787 Figure 1. (a) The unheated and heated at 400 °C H<sup>13</sup>CO-Tholins recorded with a sample  
788 spinning rate of 60 kHz and interscan delay of 30 seconds in the comparison with the IOM  
789 Murchison and Orgueil (Cody et al., 2005, 2002, Gardinier et al., 2000). Deconvolution of the  
790 solid state <sup>13</sup>C DP/MAS NMR spectra of (b) the initial H<sup>13</sup>CO-tholins (c) and the 400 °C H<sup>13</sup>CO-  
791 tholins

792 Figure 2. Raman spectrum of the 400°C H<sup>13</sup>CO- tholin. The G and D bands at ~1537 cm<sup>-1</sup> and  
793 ~ 1310 cm<sup>-1</sup>, respectively reveal a polyaromatic structure typical of coals and kerogens. The  
794 low signal-to-noise ratio and spectral artifacts at the edge are due to the high fluorescence  
795 background only partly corrected by the baseline correction.

796 Figure 3. (a) Normalized XANES spectra of the 400 °C H<sup>13</sup>CO and H<sup>12</sup>CO-tholins dominated  
797 by aliphatic (-CH<sub>x</sub>), C=C and C=O functional groups and (b) XANES and NMR methods  
798 correlation of the 400 °C H<sup>13</sup>CO-tholins with the re-produced data from Le Guillou et al. (2018)  
799 and Le Brech et al. (2016).

800 Figure 4. Corrected infrared spectra of H<sup>12</sup>CO and H<sup>13</sup>CO-tholins samples. Highlighted  
801 attributions (CH<sub>3</sub> and CH<sub>2</sub> anti-symmetric modes and C=O and C=C stretching modes) are  
802 those used for quantitative study.

803 Figure 5. Detailed components of CH<sub>x</sub> mode in different phases (liquid, crystalline, amorphous)  
804 in an example of alkane C<sub>20</sub>H<sub>42</sub> (modified from MacPhail et al., 1984).

805 Figure 6. (a) Fit of the alkyl band of an IOM extracted from the Orgueil chondrite, obtained  
806 through the artificial bees' algorithm. This fit was achieved with the 7 Gaussian model and all  
807 components FWHM forced to 35 cm<sup>-1</sup>. (b)-(c) Illusion of uncertainty of spectral decomposition  
808 methods in different scenarios: Gaussian fix a FWHM of 35 cm<sup>-1</sup> (Blue); Voigt fix a FWHM of  
809 20 – 40 cm<sup>-1</sup> (Yellow); Gaussian free a FWHM range of 20 - 40 cm<sup>-1</sup> (Green), Voigt free a  
810 FWHM range of 20 - 40 cm<sup>-1</sup> (Red).

811 Figure 7. Normalized spectral intensity of the organic functional groups derived from the  
812 calibration of C=O and C=C and CH<sub>x</sub> groups using the DP/MAS NMR and FTIR calibration  
813 for the 400 °C H<sup>13</sup>CO- tholins on (a) Murchison (CM2), (b) Orgueil (CI1), (c) EET 92042 (CR2)  
814 and (d) Tagish Lake (C2-ung) IOM in the comparison with the SP/MAS <sup>13</sup>C NMR spectra  
815 (Cody et al., 2005) and CP/MAS <sup>13</sup>C NMR (Gardinier et al., 2000).

816 Figure 8. Comparisons plots between (a) n<sub>CH<sub>x</sub></sub>/n<sub>C=C</sub> and n<sub>CH<sub>x</sub></sub>/n<sub>C=O</sub> ratios (b) n<sub>CH<sub>x</sub></sub>/n<sub>C=C</sub> and  
817 n<sub>C=O</sub>/n<sub>C=C</sub> of IOM functional C groups (Murchison (CM2), Orgueil (CI1), EET92042 (CR2)  
818 and Tagish Lake (C2-ung)) derived from FTIR-NMR quantification and the previous NMR's  
819 measurement (Cody & Alexander, 2005; Gardinier et al., 2000; Yabuta et al., 2005).

820 Figure 9. Comparisons plots between (a) H/C and O/C atomic ratios of studied coals (b)  
821 Normalized intensity of C=O, n<sub>C=O</sub> + COOH(R) vs. vitrinite reflectance, (c) n<sub>CH<sub>x</sub></sub>/n<sub>C=C</sub> vs. vitrinite  
822 reflectance and (d) Calculated A<sub>CH<sub>x</sub></sub>/A<sub>C=C</sub> vs. vitrinite reflectance derived from FTIR-NMR  
823 quantification and the previous NMR's measurement (Wei & Tang, 2018).